2015

The Psychometric Properties of the Professionalism Assessment Rating Scale

Jennifer K. Olivetti
Philadelphia College of Osteopathic Medicine, jenniferoli@pcom.edu

Philadelphia College of Osteopathic Medicine

Department of Psychology

THE PSYCHOMETRIC PROPERTIES OF THE PROFESSIONALISM ASSESSMENT RATING SCALE

Jennifer K. Olivetti

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Psychology

June 2015

PHILADELPHIA COLLEGE OF OSTEOPATHIC MEDICINE
DEPARTMENT OF PSYCHOLOGY

Dissertation Approval

This is to certify that the thesis presented to us by Jennifer K. Olivetti on the 18th day of May, 2015; in partial fulfillment of the requirements for the degree of Doctor of Psychology, has been examined and is acceptable in both scholarship and literary quality.

Committee Members' Signatures:

Robert A DiTomasso, PhD, ABPP, Chairperson

Barbara A Golden, PsyD, ABPP

Harry J Morris, DO, MPH, CPE, FACOFP

Robert A DiTomasso, PhD, ABPP, Chair, Department of Psychology

Acknowledgements

With the deepest gratitude, I would like to thank my chair, Dr. Robert DiTomasso, for his gracious guidance and his near magical ability to make me feel that I could persevere with ease. I would also like to thank Dr. Harry Morris for his ability to connect with his students and patients, and make each of us feel understood. Dr. Barbara Golden, likewise, manifests a unique ability to exude peace and equanimity in the midst of chaos. My committee members, Dr. DiTomasso, Dr. Morris, and Dr. Golden, are exemplary models of servant leaders who themselves serve as texts for their students.

I would also like to express my profuse thanks to Penny Patton for making it possible for me to experience firsthand the preparation and training that go into creating simulated encounters with SPs, for navigating the database to ensure I had the data to complete this study, and for taking the time to learn about my dreams and encourage me in my life's work. Dr. Tony Errichetti, one of the authors of the PARS, was extremely giving of his time, and he proved uniquely able to impart a sense of the passion and intent underlying the PARS. Deb Castellano, the PCOM registrar, provided invaluable help in compiling the data and also proved a continual source of caring and support for me throughout this project.

Of course, none of this would have been remotely possible without the deep familial ties that I have been so blessed to have in my life and for which I feel boundless gratitude. In particular, my mother and father, Dr. Joseph and Kathy Olivetti, supported me in ways both innumerable and constant throughout this project. My siblings; Dreon,

Jude, David, and Mark, and their families as well, proved a constant source of emotional and spiritual support, as did my most dear Aunt Mary. Linda Wright, my “coach,” also played a powerful supporting role throughout my years as a doctoral student, and beyond. I would also like to express my thanks to David Morand for helping me to “rediscover" the nascent writer within me.

Abstract

The issue of professionalism in the training of medical students has become a major focus within the field of medical education. For years, the Philadelphia College of Osteopathic Medicine (PCOM) has measured the quality of osteopathic medical students (DO students) interpersonal and communication skills through SP (SP) encounters in which SPs rate the DO students relative to eight criteria (items) which, in aggregate, are known as the Professionalism Assessment Rating Scale (PARS). These criteria are linked in the literature to patient outcomes, patient adherence, patient satisfaction, and malpractice. This study investigated the psychometric properties of the PARS, based on data from the SP encounters of 205 osteopathic medical students from a PCOM class during their first three years of medical school. Results showed the PARS to be comprised of a single, highly reliable factor. This single factor accounted for between 58.25-72.92 % of the variance across 13 separate SP encounters for which the total scale coefficient alpha ranged from .84-.95. Results of an analysis of variance indicated that performance significantly improved across each of the three years (*M* Year 1 = 6.30, *SD* = .45; *M* Year 2 = 6.40, *SD* = .44; *M* Year 3 = 6.48, *SD* = .58) and that the DO students improved consistently over the 3 years. A positive linear correlation was found between Year 1 and Year 2 (r = .46, r^2 = .22*)*, Year 2 and Year 3 (r = .42, r^2 = .18*)* and Year 1 and Year 3 (r = .36, r^2 = .13*)*. PARS scores were shown to correlate significantly with Clinical Clerkship Evaluations (CCEs) in three of the 39 SP encounters. Educational and clinical implications of these findings are discussed, and additional suggestions for future research are put forward.

Table of Contents

List of Tables

Chapter 1: Introduction

Statement of the Problem

The issue of professionalism and the quality of the physician-patient relationship in the training of medical students has in recent years become a major focus within the field of medical education. For example, the competency requirements of medical boards and professional societies increasingly reflect the vital of interpersonal dimensions that are integral to physicians' communication skills.

Of particular import and emphasis is physicians' capacity for creating and fostering a medical *working alliance* with patients (Zolnierek & DiMatteo, 2009; Hochberg et al., 2010). Working alliance refers to the agreement and collaboration between physician and patient on such elements as specific goals for treatment and the specific tasks necessary to fulfill those goals, as well as the overall emotional bond, in the form of trust, that forms between each physician and his or her patients (Bordin, 1979). In this sense of the term, working alliance references the overall quality of interpersonal communication and of the interpersonal relationship that emerges between physician and patient, over and above physicians' application of merely technical skills (Hamelin, Nikolis, Armano, Harris, & Brutus, 2012). Research shows that a positive working alliance is critically important in facilitating positive medical outcomes (Bennett, Fuertes, Keitel, & Phillips, 2011; Ferreira et al., 2013; Fuertes et al., 2007; Fuertes, Anand, Haggerty, Kestenbaum, & Rosenblum, 2014; Schoenthaler et al., 2014).

Interpersonal and communication skills, key components of professionalism that contribute to the quality of the working alliance, are undoubtedly a high priority

educational initiative in the training of medical students (Travaline, Ruchinskas, & D'Alonzo, 2005). For example, research shows that medical students who possess characteristics that contribute to the working alliance, such as empathy, are more likely to be nominated by their classmates on professionalism attributes (Hojat, Louis, Daniel et al., 2011).

Purpose of the Study

For the past several years, the Philadelphia College of Osteopathic Medicine (PCOM) has measured medical students' communication and interpersonal skills through the use of a *standardized patient* module. Standardized patients (SPs) are professional actors who play the role of patient in simulated medical exams, which are performed by medical students. SPs are thus used to simulate office visits with medical students, and subsequent to these visits the SPs rate students' performances relative to eight criteria: a) established and maintained rapport, b) demonstrated empathy, c) demonstrated confidence, d) use of appropriate body language, e) elicited information clearly and effectively, f) actively listened, g) provided timely feedback/information/counseling, and h) conducted thorough and careful physical exam or treatment. These eight criteria are, in aggregate, known as the Professionalism Assessment Rating Scale (PARS). PCOM has thus operationalized professional communication relative to the quality of the physician-patient relationship, as well as to the quality of the physical examination or treatment provided by the medical student -- captured by the PARS. The scoring rubric for the PARS is shown in Table 1.

Table 1

Professionalism Assessment Rating Scale (PARS)

Skills	Ratings								
Relationship Quality	Low-level (1-3)			Mid-level (4-6)			Superior (7-9)		
1 Established and maintained rapport *alert staff if these ratings are given to a student	1 *Reportable: So poor, student should be reported**	2 Poor	3 Mediocre	4 Fair	5 Adequate	6 Good	7 Very good	8 Excellent	9 *Rewardable: So good, the student should receive an award**
2 Demonstrated empathy	1	2	3	4	5	6	7	8	9
3 Demonstrated confidence	1	2	3	4	5	6	7	8	9
4 Used appropriate body language	1	2	3	4	5	6	7	8	9
Examination Quality	Low-level (1-3)			Mid-level (4-6)			Superior (7-9)		
5 Elicited information clearly, effectively questioning	1	2	3	4	5	6	7	8	9
6 Actively listened	1	2	3	4	5	6	7	8	9
7 Provided timely feedback/ information/ counseling	1	2	3	4	5	6	7	8	9
8 Conducted a thorough, careful physical exam or treatment	1	2	3	4	5	6	7	8	9

The developers of the PARS claim that it measures professional communication skills of doctor of osteopathy (DO) students, based on two factors: the quality of the relationship between the student physician and patient and the quality of the examination the student physician performs. The quality of relationship is proposed to be measured by establishment and maintenance of rapport, demonstrated empathy, demonstrated confidence, and appropriate body language. In contrast, the quality of the examination performed by the student physician is presumably measured by how well the student physician elicited information clearly and effectively; actively listened; provided timely feedback/information/counseling; and conducted a thorough, careful physical exam or treatment.

The PARS draws upon many constructs that are commonly found in the extant literature relative to measuring medical students' professional communication skills. However, to date, there has been no research has supported the psychometric properties of this scale. When measuring important characteristics of medical student-patient interaction, as with any other population, the use of valid and reliable measures is critical. Otherwise, what one believes is being measured may simply not be the case. In technical terms, lack of stability of a measure, or the failure for items to represent a homogeneous content domain, may threaten both the stability and internal consistency of measurements.

The purpose of the present study was to determine the reliability and validity of the PARS by evaluating whether or not the PARS actually measures the quality of the relationship the student physicians are able to establish with their patients and the quality

of the examination the student physicians conduct, as designated by the previously listed constructs. This study determined whether the assessment of these constructs is conducted in a reliable and valid fashion. This research will help to provide important answers related to the training of PCOM medical students, which, in turn, may be useful in enhancing the quality of the training offered. Details of specific items used in the construction of the PARS will be presented subsequent to the review of the literature that follows.

Literature Review

Conceptualization of Professionalism in Prior Literature

Professionalism and the nature of the physician-patient relationship have been addressed and defined in numerous ways. Ethical interpretations of professionalism view professionals as public *servants* who uphold a social contract of mutual trust with the public they serve (Cruess & Cruess, 2008; Cruess, Cruess & Steinert 2009; Cruess, Johnston & Cruess, 2004; O'Sullivan & Toohey, 2008; Stern, 2006). A developmental model describes professionalism along a continuum, ranging from the perspective of an individual to that of a complex system of interpersonal relationships (Batalden, Leach, Swing, Dreyfus, & Dreyfus, 2002; Daley, 1999; Hilton & Slotnick, 2005). Professional competence, in both ethical and developmental use of the term, includes dimensions that are cognitive in nature (obtain and utilize knowledge to solve problems), that are integrative (incorporate biopsychosocial factors in clinical reasoning), that embody a relational function (effective communication), and that tap into an affective/moral

dimension (possess affective awareness, willingness, and patience to make humane judgments; Epstein & Hundert, 2002).

From a developmental perspective, for example, osteopathic physicians take the Osteopathic Oath, during which they promise to remain loyal to their profession, to be mindful of their responsibility to preserve patients' health, to keep their patients' confidence and respect, and use good judgment in performing their duties within their skill and ability (American Osteopathic Association, Osteopathic Oath, n.d.). Service (meant here as a derivative of *servant*) is usefully thought of as the product of the interaction between two parties. Physicians and patients alike each bring their own unique expectations, personalities, biologies, and psychosocial factors into the physician-patient relationship. Physicians who are aware of their own expectations of that relationship, who are able to master their own affects, who can monitor their own biases, and who can adjust their degree of collaboration and communicate in ways that meet the expectations and needs of their patients will be more successful in fulfilling their oaths as public servants. Medical professionalism is ethically bound to the relationships within which physicians practice and the skills necessary to establish and maintain those relationships. Such skills include interpersonal and communication skills of the physician, which have been shown to influence patient outcomes (Heisler, Bouknight, Hayward, Smith, & Kerr, 2002). Thus, the development of these skills as key components of professionalism is vital to the ethical treatment of patients.

Interpersonal and communication skills are intrinsic to professionalism, and the nature of the physician-patient relationship has also been addressed and studied in the

literature through a developmental perspective. As such, professional competence is built on a foundation of knowledge, which is then combined with technical and interpersonal and communication skills, as physicians-to-be enter the context of clinical settings. For example, as medical students progress through medical school, they first acquire knowledge through reading and lectures. Then, by the time medical students are in their residencies, they combine their personal knowledge of biology, symptomatology, and decision-making formulas with the individual patient's symptomatology, biology, psychology, and socioeconomic contexts, as they begin to make real-life diagnostic decisions. Accordingly, what was once explicit background knowledge in a first or second year medical student ultimately becomes integrated with clinical experiences in a way that translates into tacit knowledge. Tacit knowledge includes the intentional use of such features as heuristics, intuition, and pattern recognition (Polanyi, 1962; Polanyi & Grene, 1969). Medical students grow in their mastery and integration of technical and interpersonal skills, until such time as they are able to diagnose patients in a timely manner, while also considering the range of contexts within which the patients must choose their course of treatment, and monitoring their own biases that may arise in the course of interacting with the patients they serve.

In short, from a developmental perspective, student physicians must ultimately learn to combine their knowledge of empirically indicated treatment, the guidance of their supervising physicians, and the motivation of their patients to make necessary behavior changes, and must accomplish all of this within the context of the healthcare system within which they practice. Clearly, the necessity of a strong physician-patient

relationship in building a successful treatment plan is one of the necessary relationships in a complex inter-relational system.

Overall, while professional competence can be thought of as the skillful application of the sum total of both technical and interpersonal skills (interpersonal skills that necessitate critical curiosity and habits of mind, such as attentiveness, self-awareness, and presence; Epstein & Hundert, 2002; Novack, Epstein & Paulenson, 1999), efforts to assess medical student competence often focus on technical skills. In so doing, the tendency is to overlook the interpersonal and communication skills that are so vital to patient outcomes and future professional behavior. While research is clear that positive patient outcomes are mediated through the physician-patient relationship, a standard measure for the quality of professional communication skills has not been established in the research. While a physician's knowledge is extremely important, the techniques he or she uses are equally important. Professionalism in communication taps the "how" of the physician's technical skill (and perhaps the "who" of the physician). The quality of a physician's interpersonal skills is itself analogous to a technique.

Prior Literature on Professional Communication and Patient Outcomes

The descriptions of physician-patient communication and the acknowledgement of the importance of such encounters exist in various genres of discussion surrounding the field of medicine, including discussions of professionalism, interpersonal skills, multicultural competence, and ethical principles. All of these are intertwined and are closely related constructs (Boulet et al., 1998; Dyche, 2007; Langenau, Pugliano, & Roberts, 2011; Ong, De Haes, Hoos, & Lammes, 1995; Van Zanten, Boulet, Norcini, &

McKinley, 2005). The general consensus is that much of the information physicians impart to patients is mediated by the quality of the relationship that is established. The emotional dimension of the physician-patient relationship, called the working alliance, is the agreement and collaboration between physician and patient on specific goals for treatment and specific tasks necessary to fulfill those goals, as well as the emotional bond, in the form of trust between physician and patient (Bordin, 1979). It is associated with patient satisfaction and adherence to treatment (Fuertes et al., 2008; Mueller, 2009). For example in a primary care sample, DiTomasso and Willard (1991) have found one key ingredient of patient satisfaction to be embedded within the perceived quality of the physician-patient relationship.

Other research indicates that patients are influenced by the relationship they have with their physician and the quality of care their physician provides during the physical examination, both of which contribute to the development of a working alliance. The association between verbal and nonverbal communication that occurs between physician and patient has long been associated with health outcomes (Beck, Daughbridge, & Sloane, 2002; Street & Buller, 1987). Poor communication between physician and patient has been linked to patient nonadherence (Eisenthal, Emery, Lazare, &Udin, 1979), and dissatisfaction for physicians and patients (Levinson, 1994), as well as to negative outcomes (Franks et al., 2005). Alternatively, strong physician communication skills (verbal and nonverbal) have been shown to contribute to the ability to establish rapport, build trust, and impart empathy and have also been linked to patient adherence, healthcare costs, patient and physician satisfaction, and to positive health outcomes

(Alexander, Hearld, Mittler, & Harvey, 2012; Epstein, Dannefer, et al., 2005; Epstein, Franks, Fiscella, et al., 2005; Epstein, Franks, Shields, et al., 2005; Epstein, Hadee, Carroll, Meldrum, & Lardner, 2007; Fuertes et al., 2007; Mueller, 2009).

Professional Communication in Medical Training

Because of the growing literature linking professional communication to patient outcomes, and in response to American Council for Graduate Medical Education's (ACGME) and the AOA's charge to assess the interpersonal and communication skills of trainees, residency programs have adopted competency-based assessment tools. Professionalism, interpersonal and communication skills, and patient care are objectively evaluated in the United States Medical Licensing Examination (USMLE), the American Board of Medical Specialties' certification, the Comprehensive Osteopathic Medical Licensing Examination- USA- Performance Evaluation (COMLEX-USA Level 2-PE), and the American College of Osteopathic Pediatricians (ACOP) program director's annual report (Langenau, et al., 2011).

The AOA's Commission on Osteopathic College Accreditation adopted seven core competencies (established by the National Board of Osteopathic Medical Examiners, NBOME, 2011; see Table 2). The AOA subsequently instructed colleges of osteopathic medicine to develop curricula in line with these competencies (AOA, 2010). Competency in "Osteopathic Philosophy and Manipulative Medicine" is cited by the AOA, in addition to the six competencies shared with the ACGME: a) patient care, b) medical knowledge, c) practice-based learning and improvement, d) interpersonal and

communication skills, e) professionalism, and 6) systems-based practice. These seven competencies are presented in Table 2.

Table 2

AOA Professional Competencies

1. Osteopathic philosophy and Osteopathic Manipulative Treatment
2. Medical knowledge
3. Patient care
a. Gathers accurate, essential information from all sources
b. Demonstrates competency in the performance of diagnostic and treatment procedures
c. Provides healthcare services that include preventative medicine and health promotion
4. Interpersonal and communication skills
a. Demonstrates effectiveness in developing appropriate doctor-patient relationships
b. Exhibits effective listening, written and oral communication skills
5. Professionalism
a. Demonstrates respect for patients/families and acts as their advocate
b. Adheres to ethical principles in the practice of medicine
c. Is sensitive to cultural diversity (ie. Religion, age, gender, sexual orientation and disabilities.
d. Is cognizant of his/her own physical and mental health in order to care effectively for patients
6. Practice based learning and improvement
7. Systems-based practice (NBOME, 2011)

Three of the seven AOA competencies have been shown to be related to the concept of professionalism and physician-patient interaction. Langenau et al. (2011) investigated the competencies as measured in the COMLEX-USA. A principal component analysis identified three factors out of the 7 AOA core competencies,

accounting for 86% of the variance between competency ratings (Langenau, et al., 2011). One factor consisted of the competencies of Patient Care, Interpersonal Skills and Communication, and Professionalism.

The interconnection of these three competencies illustrates the integration of multiple factors, all of which contribute to the working alliance and the quality of the physician's examination and treatment of the patient. As the medical profession seeks to develop and maintain the esteemed social role of its members, efforts are being made to document physician traits that positively contribute to the physician-patient relationship. The importance of professional communication in the form of the quality of the techniques and not solely the performance of technique has become the focus of medical accrediting and licensing boards.

Efforts to monitor student professionalism and their interpersonal and communication skills have shed light on changes in professionalism that occur over the course of medical education. Because students are less experienced than seasoned physicians, their ability to incorporate all of the nuances that create a strong working alliance is especially susceptible to stress (Borrell-Carrio & Epstein, 2004). Decreases in professionalism have been shown to occur during the course of medical school and can have implications for student well-being during medical school, as well as can predict future professional behavior (e.g... increased stress, censure by licensing boards, Borrell-Carrio & Epstein, 2004; Moffat, McCannachie, Ross, & Morrison, 2004; Papadakis, et al., 2005). Medical school is commonly known for being stressful (Moffat, et al., 2004). First year medical students who have been shown to mirror their peers prior to the start of

medical school show a significant increase in stress beginning with the start of classes (Aktekin et al., 2001). Medical student stress has been shown to rise significantly in the first year of medical school and to be related to academic concerns (Moffat et. al., 2004). Medical student professionalism and empathy decline as students' stress levels increase (Bellini, & Shea, 2005; Diseker & Michielutte, 1981; Hojat, et al., 2004; Shanafelt et al., 2005). This has been related to the nature of medical training (Elnicki, Lescisin, & Case, 2002; Hafferty, 2004; Silver & Glicken, 1990). Professionalism (of which communication is a part) should be monitored as an indicator of compromised student well-being.

Much research has been conducted on the relationship between empathy, student well-being and the future behavior of medical students. Empathy is significantly correlated with interpersonal and communication skills (Langenau et al., 2011). At times, empathy (an aspect that the PARS claims to contribute to the quality of the relationship developed between a physician and patient) surprisingly decreases over the course of medical school training (in allopathic schools, but not in osteopathic schools; Hojat et al., 2004, Newton, Barber, Clardy, Cleveland, & O'Sullivan, 2008; Veloski & Hojat, 2006; West & Shanafelt, 2007).

Feighny, Monaco, and Arnold (1995) found that physicians' communication skills were improved when they were trained in empathy. Increased physician empathy is linked to improved patient outcomes (in patients with diabetes Hojat, Louis, Markham, et al., 2011) and patient satisfaction (Hojat, Louis, Maxwell, et al., 2011). Thus the measurement of medical students' interpersonal and communication skills, such as

empathy, undoubtedly appears important. For example, students' level of empathy predicts subsequent performance in medical school (Hojat, Gonnella, Mangione, Nasca, Veloski, Erdmann, Callahan, & Magee, 2002). Medical students high in empathy were reported to be more likely to be nominated by their classmates on professionalism attributes (Hojat, Spandorfer, Louis, & Gonnella, 2011). Higher empathy scores have been found to predict higher competence in primary care skills in six third year clerkships (Hojat et al, 2002). Third year medical students' level of empathy predicted empathy ratings given by residency directors 3 years later (Hojat, et al., 2009). The values and behaviors of medical students in training have justifiably become a source of concern and a focus of attention in medical schools.

Medical students' professional behavior also should be monitored because it is linked to future professional behavior, which, in turn, is linked to patient behavior and outcomes. Professional misconduct in medical school has also been shown to be a predictor of future censure by state medical boards (Papadakis et al., 2005). Poor professional behavior has been shown to predict poor patient satisfaction, as previously discussed. Patients who are less satisfied with their medical care are less likely to adhere and more likely to sue (DiTomasso & Willard, 1991).

Necessarily, medical licensing boards have developed means by which to measure the interpersonal aspects of professional behavior. The NBOME and the ACGME now require interpersonal and communication skills to be assessed in objective structured clinical exams (OSCEs). These OSCEs utilize SPs who use not only checklists to determine whether or not the appropriate techniques were performed but also rating

scales that assess the quality of the communication and physical exam or treatment. Epstein and Hundert (2002) define professional competence as "the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individual and community being served" (p. 226). When assessing medical students' interpersonal communication skills, not only do the appropriate questions need to be asked and techniques applied to the patient examination or treatment, but the quality of relationship established through the manner of communication and the manner with which the physician conducts the examination and treatment must also be assessed. For example, a medical student may ask the appropriate question, "Does your family have a history of diabetes?" in a brisk manner while standing awkwardly and not looking at the patient. The nonverbal communication is also important and is best assessed through responses in actual or simulated interactions.

As discussed previously, students' performance in medical school have implications for student well-being and future behavior. Professionalism including interpersonal and communication skills should be monitored (Ainsworth & Szauter, 2006). The values that professional communication skills convey and the behaviors that communicate those values have become the focus of attention in medical schools. Medical schools use various means to monitor student competence in professionalism and interpersonal and communication skills.

The measurement of medical student professional skills. Numerous means are used to measure medical students' competence. Given the large amount of empirical

evidence showing the importance of interpersonal and communication skills to patient outcomes, attention has been given to measuring these skills in medical students. Formative and summative assessments consisting of multiple-choice exams, checklists, rating scales, direct observation, and simulations are used to measure and teach medical students professional communication skills. Medical schools use direct observation by faculty members, clinical supervisors and SPs. Educators and licensing boards seek to assess not only to assess the knowledge of how and when to use communication techniques, but also perhaps more, the quality of the interaction between the student physician and the patient. Observations by faculty members, clinical supervisors, and SPs are employed in various combinations to assess the quality of medical students' interpersonal and communication skills. However, no universal standard exists to measure student professionalism. The usefulness of any given method can be determined by five criteria: (a) the validity- if the tool measures what it claims to measure; (b) the reliability- if the tool is accurate and reproducible; (c) the impact on future learning and practice; (d) the acceptability of the measures to students and faculty; and (e) the individual, institutional, and societal cost of the measure (Van Der Vleuten, 1996).

Multiple-choice exams. The quality of medical student communication and interpersonal skills is not best measured by multiple-choice examinations. While multiple-choice tests may determine student knowledge about communication that should take place (appropriate introduction and informed consent) and appropriate wording (culturally and socially acceptable choices) to best engage patients, the affective quality and manner with which student physicians conduct the physical part of their patient

exams cannot be replicated in a multiple-choice exam. The nature of a student physician's communication (verbal and nonverbal) in the context of the physical examination or treatment is more accurately observed through behavioral determinants. While highly reliable because of the large number of items that can be assessed in a short amount of time, in multiple-choice examinations suffer in validity because they do not assess the manner in which student physicians communicate. Students, faculty, and the larger society do benefit from the use of multiple-choice tests, however, because of their ability to show whether or not students have the knowledge base to move forward in their education. They can be given in a fair, confidential, and inexpensive manner to a large number of students simultaneously. They cannot, however, assess the quality of the working alliance that a student is able to form with his or her patients or the ability of a student to actively listen to a patient and collaborate with that patient on a treatment plan. Given the empirical basis of the impact of physicians' interpersonal and communication skills to patient outcomes, for medicine as a profession, in service to the public should refrain from using multiple choice examinations as the sole measure of medical student competence.

Checklists. Checklists consist of lists of skills for which an observer notes the behaviors completed by the student. Colleges have individually developed formative assessment tools to determine whether or not competency goals are met. For example, The Kirksville (MO) College of Osteopathic Medicine-A. T. Still University developed a Matrix for Quality Enhancement (MQE)--a checklist-based curriculum assessment tool. In an effort to integrate the seven core competencies and the benchmarks of the

COMLEX-USA Level 2-PE with the knowledge based learning of the first-year osteopathic medical students' medical microbiology class, Lockwood Tucker-Potter, and Sargentini (2009) implemented the MQE using SPs) in 16 encounters. SPs provided feedback on professional communication skills.

Rating scales. Global rating scales, which rate characteristic overall performance, are typically given, along with verbal comments, at the end of clerkship rotations. While these measures do address the quality of physician-patient relationships, they are subjective, and standards may not be clearly articulated, bringing into question the validity of such measures. Also, they are given much too infrequently to establish good reliability in such tools (Pulito, Donnelly, & Plymale, 2007; Pulito, Donnelly, Plymale, & Mentzer, 2006). Still, direct assessment of patient care and these ratings scales have long been used to rate medical students' interpersonal and communication skills, including the manner in which they conduct the physical examinations. These tools provide a means of summative assessment for professional competence at the end of each clerkship that is relatively simple to complete for faculty and allows students to compare their results to those of other students and to their own future and subsequent performance on the rating scales.

Direct observation. In an attempt to increase the reliability of supervisor ratings, assessments that allow for a number of direct observations were developed (e.g., the mini-clinical-evaluation exercise [mini-CEX], Norcini, Blank, Duffy, & Fortna, 2003; and the long case, Norman, 2002). In direct observations, supervising physicians observe medical students conducting physical examinations and focused history taking with

actual patients. After a 10- to 20-minute patient encounter, the resident presents a diagnosis and treatment plan to the faculty member, who rates the student and may provide formative feedback. Faculty members may also review videos of patient encounters and provide feedback in much the same manner (Epstein, Dannefer, et al., 2004). Direct observations are often coupled with oral presentations, literature reviews, and assessments of clinical reasoning. These assessments with live patients provide a greater number of presenting problems, physical findings, and clinical settings (Cox, Irby, & Epstein, 2007), and they can be as reliable as encounters with SPs (Norcini, Blank, Duffy, & Fortna, 2003; Norcini & McKinley, 2007; Van der Vleuten, Norman, De Graaf, 1991; Wass, Jones, &Van der Vleuten, 2001).

Simulations. Simulations take many forms; multiple-choice, computer-based simulations, part-task trainers, integrated simulators, and SP encounters—all built to reproduce the clinical environment relevant to the task being assessed. They may include specific aspects of a given task, or an authentic examination room with SPs. The degree to which a simulation mimics real life depends upon the aspects of the simulation that are being measured and the degree to which various aspects of the environment are important to reproduce (Boulet & Swanson, 2004)

Multiple-choice simulations. Multiple-choice examinations are considered low-fidelity simulations which assess decision-making skills. The degree to which they simulate the decision making that occurs in real patient encounters depends on the detail given in the patient description (Swanson & Case, 1993).

Computer-based simulations. Computer-based simulations give medical students a brief description and history of a patient and allow them to choose diagnostic techniques, order therapeutic interventions, and monitor patient progress. Risk avoidance, timeliness, effectiveness and thoroughness are assessed using algorithms created by patient care strategies of experienced physicians. These simulations are shown to be more precise and valid than multiple-choice simulations (Clauser, Margolis, & Swanson, 2002; Dillon, Clyman, Clauser & Margolis, 2002), but they can be expensive to administer and require a significant amount of human and financial resources to create and score (Boulet & Swanson, 2004). While they do simulate patient encounters and require medical students to consider biopsychosocial factors in the care of their patient, they do not specifically or directly assess interpersonal or communication skills, much less the quality of those skills.

Part-task-trainer simulations. Part-task-trainers replicate only pieces of the clinical environment, such as specific parts of the body and can be used to train medical students in the motor skills necessary to perform maneuvers such as pelvic exams, joint injections, suturing, catheterization, or incisions. While it may be argued that the quality with which the student performs these maneuvers can imply the nonverbal communication of empathy that occurs when a physician is careful with his or her patients, the purpose of these simulations is not to assess interpersonal skills.

Integrated simulations. Integrated simulators connect mannequins to computers that determine the symptoms they portray (Schwid, 2002). They talk and have detectable pulses. Their pupils dilate. Some even urinate. Computers can cause monitoring

equipment to show responses as simple as an electrocardiogram and as complex as intracranial pressure. Interventions show reactions such as those after oxygen has been administered. High-fidelity integrated simulators are expensive and require technology that must be maintained. Still, they can be used to train students to respond to situations for which errors cannot be reversed, as well as those that cannot be enacted by SPs (emergency situations involving trauma). This provides crucial training for students before they reach residencies and are faced with live patients. In this way these simulations serve as a safety measure for the public. Students can practice and gain confidence, with less demand on faculty. Integrated simulators, like part-task-trainers, lack validity in their assessment of interpersonal and communication skills, because the integrated simulators are simply not human.

SP simulations. SPs are trained actors who enact specific patient roles consistently in SP encounters with medical students. Observations of these interactions are made by faculty and supervising physicians. These observations are combined with other learning opportunities, just as the direct faculty observations previously discussed. In addition, SPs can provide their own ratings, from the perspective of the patient, in the form of checklists, verbal feedback, and global rating scales. While checklists can assess interpersonal and communication skills along with technical skills, global rating scales enable SPs to capture the quality of the interactions. For example, a checklist item requires the rater to record whether or not the student performed a particular skill, such as "asked for permission to touch patient." A global rating scale requires the rater to evaluate "how well" or "the quality with which" a student palpated the patient's

abdomen. Verbal feedback given by the SP can address the quality with which the student conducted both the technical and interpersonal skills during the encounter. SPs are trained to use the scales, and interrater reliability is established to ensure the measurements taken in these encounters are reliable (Shirazi et al., 2014). Although encounters with actual patients have been shown to be equally as reliable as those with SPs (Norman, 2002; Van der Vleuten, Norman, & DeGraaf, 1991), the use of SPs allows institutions of medical education to control the number and type of interactions for which students are assessed. SPs can also be trained for better control in the observation of specific communication skills, such as body language or interviewing skills (Imel, et al., 2014) and therefore provide more targeted feedback to students as well. The use of planned SP scenarios ensures that the scope and volume of assessments provide a reliable valid measure of medical students' abilities.

Objective structured clinical examinations (OSCEs). Medical schools and licensing boards utilize SPs to complete checklists and global rating scales in OSCEs, which incorporate a series of time-limited SP encounters in summative assessments of medical students and candidates for licensure. OSCEs are part of the USMLE and the COMLEX-USA, which medical students must take before they can proceed to residency. The SPs or faculty members complete checklists or global rating scales to evaluate medical students' competence. SPs can help evaluators complete an adequate number of examinations, which include the assessment, working alliance and clinical professionalism within a limited time frame in order to ensure the examinations are reliable.

SP encounters can be targeted to address specific formative goals. Educators who seek to develop the professional competence of their students can instruct SPs to address the working alliance in their feedback to the student. Likewise, faculty observers of SP encounters can offer additional feedback regarding the physician-patient relationship (Epstein, et al., 2007). Feedback can be tailored to the medical student's appropriate level of development.

The numerous means developed to measure medical students' professionalism show its importance. Monitoring of student development of professionalism is crucial and remediation can have implications for the students' subsequent training, clinical effectiveness, and future career.

The measurement of medical student professional competence according to the PARS. Having thus reviewed the literature on professional competence in medical settings, and having reviewed the various manners in which such competencies have been measured in the extant literature, a detailed analysis of the measurement characteristics of the PARS instrument is conducted, as this instrument is the major focus of the present study.

Quality of physician-patient relationship. Competent communication skills are key to building the physician-patient relationship and can greatly influence patient behavior and medical outcomes. The PARS conceptualizes the quality of the student physician's ability to build a physician-patient relationship as an aggregate of the quality with which student physicians (a) establish and maintain rapport, (b) demonstrate empathy, (c) demonstrate confidence, and (d) use appropriate body language.

Research shows that physicians who communicate in ways that build rapport with their patients will accrue benefits in the form of earned patient trust (Epstein et al., 2007). Patients' trust is an important variable relative to their choice of, and response to, physicians. Physicians can establish rapport and earn trust when they give clear expectations for the office visit, such as by telling the patient how much time they have together and what they intend to do in that time, while at the same time inviting the patient to collaborate with them on the course of treatment. Research shows that when considering development of trust in surgeons, patients value and respect autonomy and verbal communication the most, followed secondarily by technical skill (Hamelin, et al., 2012). Thom, Kravitz, Bell, Krupat, & Azari (2002) surveyed patients and found that those with low trust levels toward their physicians were more likely to report that a service they asked for was not provided, less likely to report symptom improvement, and less likely to have the intention to follow their physician's advice.

Physicians with the interpersonal skills to be mindful of their verbal and nonverbal communication with their patients can influence their patients' behavior in ways that will improve patient satisfaction and adherence to treatment. The manner in which the physicians present themselves can contribute to their patients' confidence in them. Confident physicians treat patients with respect, and they are not afraid to tell a patient when a question or problem falls outside their area of expertise, nor do they fear, in turn, asking for help from a supervising physician or colleague. Such confident physicians possess the ability to maintain the presence of mind such that they can consider their own emotional states, while at the same time observe patients' affective

cues relative to their presenting problems (Epstein, Alper, & Quill, 2004). Physicians who ask questions and provide information based on patients' communication, both verbal and nonverbal, instill confidence in their abilities with their patients. Alexander et al. (2012) found that patients who perceived higher quality in their physician patient communication had higher levels of activation. Activation, the confidence and ability to inform oneself about one's health related behavior and to communicate with the physician relative to concerns about one's treatment, is important because adherence to physician-suggested behaviors is more likely in patients who are more activated (Becker and Roblin, 2008). For example, a physician may ask a patient how she thinks her increased workload is affecting her sleep, but if that physician does not make eye contact and pause long enough for the patient to thoroughly answer the question, that physician may not build the rapport necessary for the best treatment of that patient.

Patients invited to collaborate in the building of their treatment plan are more likely to be more satisfied with their resultant treatment if their expectations for participation in their own healthcare match those of their physician (Jahng, Martin, Golin, & DiMatteo, 2005). Patients are more likely to follow their physician's prompts in regard to their role in their own treatment (Berry et al., 2008; Blanquicett, Amsbary, Mills, & Powell, 2007). Further, fair treatment, regardless of socioeconomic background, is a motivator for patients to mirror their physicians' level of engagement (Berry, et al, 2008). Thus, physicians who interact with their patients in a manner that solicits patient engagement in their own diagnosis and treatment will witness a higher level of patient

activation and therefore witness the higher level of patient adherence with that engagement. Patients who are encouraged to share their health beliefs about specific treatments will collaborate with the physician to devise plans that the patient is more likely to follow. For example, a patient who comes to the office and expresses that she is there for antibiotics in order to ensure she will not miss work is highly likely to adhere to a treatment of antibiotics. Alternatively, a patient who is encouraged with attentive posture, time to answer, and an encouraging head nod will be more likely to share her belief that antibiotics should not be used unless as a last resort. This patient is much less likely to adhere to a treatment of antibiotics but may welcome an alternative treatment, such as an osteopathic manipulation or a nasal rinse. The latter patient is also likely to be much more satisfied with, and trusting of her physician because he or she did listen to her, having been shown empathy by her physician, she may actually feel better leaving the office.

The ability of physicians to ask questions about their patients' needs and to attend to cues their patients give about their desires and expectations for treatment allows physicians to build a strong working alliance in which patients are more satisfied. Research shows patients are more satisfied when their expectations for the length of time spent in their office visits are realized (Lin, et al., 2001), and when their expectations for participation in their own healthcare match the expectations of their physicians (Jahng, Martin, Golin & DiMatteo, 2005). When patients feel their need for information has been met, they are more satisfied and report fewer symptoms (Kravitz et al., 2002).

Empathy has long been considered a foundation of the physician-patient relationship (DiMatteo, 1979; Levinson, 1994) as well as a factor in patient outcomes (Hojat, Gonnella, Mangione, Nasca, & Magee, 2003; Hudson, 1993). Effective physicians empathize with their patients' needs (Blasi, Harkness, Ernst, Georgiou, & Kleijnen, 2001). Effective communication (Heisler et al., 2002) and physician empathy levels have been linked to positive clinical outcomes in patients with diabetes (Hojat, Louis, Markham, et al., 2011). Patients are more likely to be influenced by physicians who are perceived as warm, empathic, and caring,- a concept referred to as referent power (DiMatteo, 1979; DiMatteo, Haskard-Zolnierek, & Martin, 2012).

Empathy can be demonstrated in many ways, including the manner in which the physician touches the patient while doing the physical examination, the respect for the patient's modesty in the way the physician drapes the patient during the examination, and the verbal and nonverbal expression of compassion for the pain a patient may be experiencing. Empathy in the working alliance has been studied extensively as a key component of professional competence. Hojat et al. (2003) have created a valid and reliable scale measuring empathy as an important predictor of professional communication as previously discussed. Empathy assists the physician in establishing a good relationship with his or her patient. The PARS defines empathy as a distinct item contributing to the quality of the relationship established during a given patient encounter. It can also contribute to the quality of examination, as it is demonstrated through a "felt-sense" in the quality with which the physician conducts the physical examination of the patient.

***Quality of examination*.** The manner in which the physician conducts the physical examination communicates much to the patient. The PARS conceptualizes the quality of the student physician's examination as a construct of the quality with which student physicians (a) elicit information (clearly and effectively), (b) listen and respond appropriately (actively listen), (c) provide timely information and counsel their patients, and (d) provide a (thorough and careful) examination and treatment of their patients. Student physicians who conduct the patient examination with professional quality communicate empathy, build rapport, and instill confidence in the patient's care. Physicians must possess strong interpersonal and communication skills because healthcare is currently much different from the healthcare of the past. While some patient-physician relationships remain traditionally passive as the patient accepts the advice of the more educated and usually of higher-social-status physician, the trend is for patients to take a much more active part in their healthcare.

Physicians who elicit information clearly from patients can bridge gaps that may occur as a result of cultural, educational, and socioeconomic differences, gaps that may interfere with the patients' understanding of the information their physician imparts (Berry et al., 2008). For example, physicians who are mindful of a given patient's lack of numerical literacy should be careful to explain treatment options in terms of relative gains and risks of vacuous procedures or treatments, in contrast to communication with patients who understand numbers and percentages (Gordon-Lubitz, 2003). One study examined internal medicine residents' listening skills and their attending physicians with their English-speaking, primarily low-income, and ethnic minority patients in a

community-based ambulatory clinic (Dyche & Swiderski, 2005). The percentage of patient problems that the physician correctly identified was 84.6 % when the patient was allowed to complete his/her agenda, 82.4% (P = .83) when the patient was interrupted, and only 59.2% when the physician did not solicit an agenda from the patient, representing a 24% drop in physician understanding. Poor listening skills decrease physicians' abilities to elicit information from their patients. Physicians who make efforts to listen to their patients' life situations and concerns are less likely to miss diagnostic information or to prescribe treatments to which patients will be unable to adhere, and are more likely to prevent unnecessary patient suffering as well as overutilization of resources (Sandella, Roberts, & Langenau, 2011).

Strong listening skills can improve the physician's ability to elicit information and provide a bridge to other forms of care. Similarly, the manner in which some physicians conduct their exams imparts empathy and can also provide a bridge to other forms of care. Physicians who allow patients to express their emotions and who acknowledge the emotions are less often viewed as uncaring (Suchman, Markakis, Beckman, & Frankel, 1997). Physicians who take time to listen to their patients initially are likely able to gather important data that might otherwise be missed. Marvel, Epstein, Flowers, and Beckman and colleagues (1999) conducted a study that found physicians interrupt and redirect patients after an average of only 23 seconds into the patients' descriptions of their presenting problems. Researchers suggest that this quick redirection results in missed opportunities for data gathering, as well as in the patient presenting concerns later in the treatment. Open communication between physician and patient about stress from

the outset of treatment helps to reduce the stigma and anxiety that may accompany referrals to psychological intervention, strengthens the working alliance, and improves compliance with treatment (Janeway, 2009). Physicians who take time to listen to their patients can avoid finding out key information only late in the treatment, information that could influence the treatment adherence and outcomes. Moreover, they are better able to discern when patients require specialty care.

The thoroughness and care taken when a physician conducts the physical examination influences patients' responses to the physician's examination. Epstein, Franks, Fiscella, et al. (2005) operationally defined patient centered care (PCC) as follows:

> Eliciting and understanding the patient's perspective—concerns, ideas, expectations, needs, feelings and functioning; understanding the patient within his or her unique psychosocial context; reaching a shared understanding of the problem and its treatment with the patient that is concordant with the patient's values; and helping patients to share power and responsibility by involving them in choices to the degree that they wish. (p. 1517)

They conducted a study that showed PCC to be a mediator between office visit duration and diagnostic testing cost, thus showing that physicians with good active-listening skills can prevent unnecessary healthcare costs.

Another aspect of the quality of the examination, one of which affects the patient's response to his or her physician, is the manner in which a physician gives a patient feedback. Patients who feel their need for information has been met are more

satisfied and report fewer symptoms (Bell, Kravitz, Thom, Krupat, & Azari, 2002: Thom, et al., 2002). The timing of the delivery of feedback, the amount and type of information given, and the manner in which the information is given have been shown to be conducive to patient outcomes. One study had participants complete a series of surveys at different intervals prior to and following office visits for acute conditions. The study concluded that physicians can improve patient understanding and reduce patient emotional distress by matching the level of information they give to the patient's level of desire for information (Miller, Brody, & Summerton, 1988). Also, when physicians must give bad news to patients and their families a silent pause after giving the news can allow the recipients to more fully comprehend the information given (Ambuel & Mazzone, 2001; Sweeny, Shepperd, & Han, 2013; VandeKieft, 2001). Physicians who are mindful of their own reactions and manner of communication can better communicate, provide higher quality treatment for their patients, and ensure the patients receive the most optimal care for each unique situation.

Summary

In training medical students, the quality of the physician-patient relationship warrants careful attention. Professionalism and the quality of the medical alliance are important in facilitating positive outcomes. More research is needed to establish the quality of measures used to assess these important dimensions. Given the important association between the quality of the physician-patient working alliance and health outcomes, the psychometric properties of scales used to measure the quality of students' professional communication skills, which contribute to that alliance, are paramount.

Otherwise medicals schools may believe they are measuring the professional communication skills of their students, when such is not necessarily the case.

Chapter 2: Research Questions and Hypotheses

1. What is the factor structure of the Professionalism Assessment Rating Scale (PARS)? The PARS was developed by a multidisciplinary team at the Philadelphia College of Osteopathic Medicine (PCOM). This team conceptualized the PARS in overall terms to measure the quality of the physician-patient interaction, but composed of two subdomains: in regard to the first subdomain, the relationship, and the second, the physical examination. In accord with the scale's author, this study proposes a two-factor structure.

Hypothesis 1: There will be a two-factor structure of the PARS, specifically quality of the exam and quality of the relationship.

2. What is the coefficient alpha reliability of the PARS scale? Patients' comfort levels with their physicians are decided based not on what their physicians say and do, but on how their physicians communicate and interact with them (Jackson, Chamberlin, & Kroenke, 2001). The PARS aims to measure this "how" in terms of the quality of the physician-patient interaction as highlighted by eight items, all of which are linked to patient outcomes (Bennett et al., 2011; Ferreira et al., 2013; Fuertes et al., 2007; Fuertes et al., 2014; Schoenthaler et al., 2014), satisfaction (DiTomasso & Willard, 1991), adherence (Mueller, 2009), malpractice (Adamson, Tschann, Gullion, & Oppenberg, 1989) and each other. It follows that the items that comprise the PARS are actually more likely to measure a single construct.

Hypothesis 2a: The PARS will demonstrate a coefficient alpha of at least .70.

Hypothesis 2b: The subscales of the PARS will demonstrate an overall coefficient alpha of at least .70.

3. Does the DO students' professionalism, as measured by the PARS, increase from the first to the second year and the second to the third year? Training has been shown to improve interpersonal and communication skills when experience alone does not improve these skills (Carvalho et al., 2014). For example, after completing communication skills training, individuals' self-ratings of clinical competence improve, as do observations of their confidence (one of the eight items on the PARS) by SPs and faculty member observers (Carvalho et al., 2014; Delvaux et al., 2004). Likewise, patients report a higher level of perceived sense of care from their physicians with training in communication skills (Moore, Rivera, Mercado, Artigues, Grez, & Lawrie, 2013). Hojat, Mangione, Nasca, Gonnella, and Magee (2005) showed that empathy ratings remained steady throughout residency programs when faculty reports at the end of residency were compared to self-report measures of empathy prior to entering residency programs. Consequently, monitoring PARS scores over the course of training is worthwhile to discern whether training programs are efficacious in improving the level of quality to which student physician are affording their patient interactions.

Hypothesis 3: Mean PARS scores will increase between the first and second years and between the second and third years.

4. What is the relationship (correlation) between the first-, second-, and third-year PARS scores? A correlation in aptitude across years of osteopathic medical training will support Hypothesis 3 and confirm whether or not training in the quality of physician-

patient interaction, as measured and taught using the PARS, is assisting in the development of these skills in the DO students at PCOM. Research shows that training, and not experience alone, improves such skills as previously discussed above in support of Hypothesis 3.

Hypothesis 4: There will be a positive correlation between first-, second-, and third-year PARS scores.

5. What is the relationship between PARS scores and ratings by clerkship supervisors? Because the Clinical Competency Evaluations (CCEs) encompass the osteopathic competencies and three of those competencies are patient-care, professionalism, and interpersonal and communication skills (all of which relate to professionalism as measured by the PARS), and because professional competence is linked to elements of the PARS, such as empathy (Hojat et al., 2002).

Hypothesis 5: There will be a positive correlation between DO students' mean (overall) PARS scores and the mean of the DO students' clerkship supervisors' ratings.

Chapter 3: Method

Overview

This archival study is based on information collected at the Clinical Learning and Assessment Center (CLAC) at the Philadelphia College of Osteopathic Medicine (PCOM). The clinical professionalism of osteopathic medical students (DO students) in the first, second and third years of the medical curriculum are assessed using the Professional Assessment Rating Scale (PARS). DO students are also assessed using Clinical Competency Evaluations (CCEs) completed by clerkship supervisors in the DO students' third and fourth years of osteopathic medical education. Approximately 2,665 administrations of the PARS were examined for this investigation. These ratings were collected from all PARS completed for all DO students in one graduating class. Each student completed a total of 13 standardized patient (SP) encounters in his or her first 3 years of medical school for which PARS are completed by SPs. Likewise, 1,435 CCE ratings from were examined for this investigation. These ratings were collected from 7 key clerkships that each student completes in his or her third and fourth years of medical school.

SPs and the PARS

The standardized patients (SPs) complete a standardized training program prior to the beginning of each term, in which efforts are made to standardize the ratings among the SPs. The 3-hour-long standardized training, facilitated by a training director with several years of experience, is complete by seven or eight SPs at a time. All of the SPs are introduced to the PARS and familiarized with the behavioral anchors used to score

the PARS. They then watch a series of 14-minute encounter videos. The SPs rate the quality of the eight items in the PARS independent of the technical accuracy of the skills completed. SPs rate the students first on whether their skills were low-level, mid-level, or superior for each item, then on the degree to which they met that level. Following each video, the SPs complete the PARS and place their scores for all items on a board so all the other SPs can see it. The SPs then discuss the scores with trainers and reconcile any divergent scoring that may have occurred. This process allows them to self-regulate for stringency/leniency and aims to standardize the ratings and achieve interrater reliability. The PARS is completed as part of a SP encounter for which the SPs act as patients and then complete a checklist of skills and the PARS for each student. A case script is used for each encounter, based upon the curriculum the students are covering in their classes at the time. For example, if they are studying the ear, nose and throat; the SP is likely to be presenting a scenario in which he or she has a head cold. SPs rate each item, using the PARS following each encounter, while the student sits in a waiting area lined with desks to write their subjective, objective, assessment, plan (SOAP) notes. When the SP finishes completing a clinical skills checklist and the PARS, he or she gives feedback to the student. The case scenario for each encounter, the length of time given for each encounter, and type of feedback (written, verbal, checklist, PARS, video) are consistent and scheduled to match the curricululm for the term in which they take place. The final encounter for each year is an Objective Structured Clinical Exam (OSCE) in which students are graded as part of the DO curriculum. The OSCE in the third year of training consists of three, 14-minute encounters, rather than one 50-minute encounter.

Clinical Competency Evaluation

PCOM requires a CCE to be completed by clerkship supervisors for each student at the end of each rotation. Each PCOM student complete a total of 23 clerkships in his or her third and fourth years of osteopathic medical school. Clerkship supervisors complete the CCE and review summative and formative comments with students prior to the end of each rotation. Clerkship supervisor ratings were collected for the same students as were the PARS ratings. The CCE is an 18-item scale used to rate DO students on the seven osteopathic core competencies. The measure is designed to include seven factors, one for each of the seven core osteopathic competencies; *Patient Care, Medical Knowledge, Practices-based Learning and Improvement, Interpersonal and Communication Skills, Professionalism, System-based Practice, and Osteopathic Principles and Practice*. The 18 items on the CCE include 1) History Taking and 2) Physical Examination designed to measure *Patient Care;* 3) Fund of Knowledge and 4) Problem Solving designed to measure *Medical Knowledge*; 5) Integration of Instruction and 6) Efficiency and Effectiveness designed to measure *Practices-based Learning and Improvement;* 7) Humanism and Interpersonal Skills, 8) Oral Presentations, and 9) Written work; designed to measure *Interpersonal and Communication Skills;* 10) Skills in Dealing with Diversity and Cultural Differences, 11) Feedback/Constructive Criticism, and 12) Commitment designed to measure *Professionalism;* 13) Collaborative Practice Skills, 14) Disease Prevention/Routine Health Maintenance, and 15) Cost Consciousness designed to measure *System-based Practice;* and 16) Osteopathic Philosophy, 17) Osteopathic Structural Exam, and 18) Osteopathic Manipulative Technique designed to

measure *Osteopathic Principles and Practice.* Each item is rated by clerkship supervisors on a 10-point, Likert-type scale, with ratings of 1 reflecting *substandard* performance, 2-5 reflecting *marginal* to *adequate* performance, 6-8 reflecting *competent* to *proficient* performance, and 9 and 10 reflecting *outstanding* performance. Students receive a recommended grade by their clerkship supervisor of Fail, Pass, or Honors Pass. Scores are collected in the Clinical Education department of PCOM and recorded on DO student transcripts.

Research Design and Design Justification

The design of this study was a psychometric analysis of the PARS, the correlation between the PARS scores and other criteria, and as a within subjects design of change in PARS scores over time. To test Hypothesis 1, a Varimax rotated principal component analysis using Kaiser's criterion (an eigenvalue of 1 as a basis for factor extraction) was conducted. To test Hypothesis 2, a coefficient alpha reliability of the total PARS and a coefficient alpha reliability of any subscales identified within the PARS was planned. To test Hypothesis 3, a one-way repeated measures analysis of variance (ANOVA) was planned to test the changes in PARS scores from first year to third year. To test Hypothesis 4, a Pearson correlational analysis was planned to identify relationships between mean PARS scores for the first-year and mean PARS scores for the second-year administrations, the mean PARS scores for the second-year administrations and the mean PARS scores for the third-year administrations, and the mean PARS scores for the first-year administrations and the mean PARS scores for the third-year administrations. To

test Hypothesis 5, a Pearson correlational analysis was planned to assess PARS scores and CCE ratings by clerkship supervisors for seven selected clerkships.

Participants

Study participants included 205 osteopathic medical students in the PCOM DO program who were rated by SPs using the PARS twice each term in their first year, twice in the first and second term and once in the third term of their second year, and three times at the end of their family medicine rotation during their third year. All members of one entering class provided useable data.

Inclusion and exclusion criteria

Students included in the study are all students who are members of one entering class and who participated in clerkships. Students who did not participate in clerkship were excluded from the study.

Recruitment

Participants were part of a large database maintained by the CLAC. All students were required to complete the designated SP encounters and receive ratings by clerkship supervisors within each year of their education. Archival data from these encounters where analyzed.

Measure

PARS

The PARS is an eight item scale that was designed to assess the professionalism of medical students. This measure was designed to include two factors; one related to *relationship quality* and another related to *examination quality*. The eight-items on the

PARS include: (a) established and maintained rapport, (b) demonstrated empathy, (c)demonstrated confidence, (d) used appropriate body language, (e) elicited information clearly, effectively, (f) actively listened, (g) provided timely feedback/information/counseling, and (h) conducted a thorough, careful physical exam or treatment. Items 1-4 were designed to measure relationship quality, and Items 5-8 were designed to measure examination quality. Item is rated (by SPs) on a 9-point Likert-type scale ranging from "how poorly" to "how well" the student relates to and examines the SP, with ratings of 1-3 reflecting low level quality (1 = so poor the student should be reported 2 = poor, 3 = mediocre), 4-6 reflecting mid-level quality (4 = fair, 5 = adequate, 6 = good), and 7-9 reflecting superior quality (7 = very good, 8 = excellent, 9 = so superior the student should be rewarded; see Table A). The scale is completed by the trained SPs upon completion of each encounter with each student. Students are scored on each item individually.

Relationship quality. Relationship quality is reflective of how a physician/medical student establishes and maintains rapport, demonstrates empathy, demonstrates confidence, and uses appropriate body language. Items 1-4 of the PARS measure relationship quality.

Establishes and maintains rapport. A medical student establishes and maintains rapport when he or she achieves and retains a positive, respectful collaborative working relationship with the patient. This rapport is as demonstrated when the medical student addresses the patient, respects the patient's modesty, obtains consent, respects the patient,

collaborates with the patient, uses appropriate body language, and exhibits appropriate professional dress and hygiene.

Demonstrates empathy. A medical student demonstrates empathy when he or she achieves (verbally or non-verbally) understanding, and shows concern and interest in the patient's medical problem and life situation through expressions of interest in the patient's life and health; expressions of concern about the patient's life and health; expressions of understanding about the effect of the physical exam; a warm and caring demeanor; and genuine expressions of interest, concern, understanding, warmth, and caring.

Demonstrated confidence. A medical student demonstrates confidence when he or she acknowledges his or her limitations without expressing them in a way that undermines the patient's confidence in him or her. This skill is exhibited when the medical student conveys a confident demeanor and offers to get help which he or she cannot provide.

Uses appropriate body language. A medical student uses appropriate body language through his or her body cues, signs, and gestures; through appropriate eye contact, and through the maintenance of a comfortable "personal distance" and "personal closeness."

Examination quality. Examination quality is reflective of how a physician/medical student elicits information, listens, gives feedback/information/counseling, and conducts a physical exam or treatment. Items 5-8 measure examination quality.

Elicits information clearly and effectively. A medical student elicits information clearly and effectively when he or she asks questions in an articulate, understandable, straightforward manner through a balance of open-ended and close-ended questions and avoidance of leading questions.

Actively listens. A medical student actively listens when he or she listens to the patient, responds to the patient's statements, and asks appropriate questions of the patient. The student makes eye contact, listens without interrupting, appears to listen while taking notes, listens, and then responds to what a patient says, and listens attentively.

Provides timely feedback/information/counseling. A medical student provides timely feedback, information, and counseling when he or she explains, or summarizes information (e.g., results of physical exams, patient education activities) or provides counseling in a clear and timely manner. This skill is exhibited when a medical student discusses the agenda, explains procedures; provides meaningful feedback, provides a timely summary, provides counseling, avoids jargon, and provides closure.

Conducts thorough, careful physical exam or treatment. A medical student conducts a thorough, careful physical exam or treatment when he or she completes the physical exam in a thorough, careful manner versus a tentative or superficial manner. This skill is exhibited when a medical student conducts a careful and smooth examination, inspects the patient, and uses the time allotted.

Procedure

Student scores for each of the eight domains in the PARS for each of the 13 separate PARS scores from SP encounters and CCE scores were extrapolated from the

Arcadia database used for recording SP encounters in the CLAC at PCOM. The data for this study had been stored in Arcadia—a database managed by the Director of the CLAC—and were extrapolated and de-identified. These data were downloaded into an SPSS datafile and cleaned for subsequent analysis. These data are part of the CLAC's SP Training Program, which is described in more detail later. The CLAC protocol includes development of cases, training of SPs, scoring of student performance, and feedback**.** Data were analyzed for PARS scores and CCEs completed for DO students of one graduating class in the CLAC. The data were imported into SPSS for statistical analysis, including factor analysis, reliability (coefficient alpha) analysis, correlational analyses, and tests of between-groups effects for demographic variables. PARS scores were analyzed for significant changes over time. Correlations between first, second, and third years were investigated. The correlations between the mean of the DO students' clerkship supervisors' ratings were determined.

Chapter 4: Results

A total of 205 Doctor of Osteopathy (DO) students, representing one PCOM class, completed 13 SP Encounters during which the SPs used the Professionalism Assessment Rating Scale (PARS) to evaluate each student. The clerkship supervisor ratings for these same DO students from seven selected clerkship rotations were also used to evaluate the psychometric properties of the PARS. The aforementioned hypotheses were expected to explain the PARS in terms of factor structure, internal reliability, change in scores across years, correlation in PARS scores across years, and the relationship of PARS scores to clerkship ratings for respective years of education.

Factor Structure

Hypothesis 1 proposed that there would be a two-factor structure to the PARS, specifically quality of the exam and quality of relationship. To test this hypothesis, a Varimax rotated principal components factor analysis, with a Kaiser criterion set to an eigenvalue of 1, was conducted to determine if the eight items measured in the PARS separated into two factors. Each study participant took part in a total of 13 separate SP encounters. Consequently, in order to determine whether or not the factor structure remained invariant across those encounters, a separate factor analysis was performed for each of the 13 SP encounters.

The results show that for 12 of the 13 SP encounters there was only one component was extracted, based on an eigenvalue of over 1. In each of these 13 cases, the single extracted factor ranged from explaining 57.31 to 72.92% of the variance. These results clearly indicate a single-factor solution substantiated repeatedly over many

encounters. The analysis thus revealed that the eight items of the PARS comprise one overall factor, not two as hypothesized by the original authors of the scale. This single factor accounts for 68.84 % of the cumulative variance of the PARS scores from the first SP encounter (Headache), 70.38 % from the second SP encounter (Eye Exam), 72.92 % from the third SP encounter (Earache/ Sore Throat), 63.09 % from the fourth SP encounter (Chest Pain/ Shortness of Breath), 61.59 % from the fifth SP encounter (Abdominal Pain/ Nausea) , 58.25 % from the sixth SP encounter (Diverticulitis/ GERD), 63.84 % from the eighth SP encounter (Angina/ Abo.), 64.69 % from the ninth SP encounter (Hip/Knee Pain, Counseling, 57.31 % from the tenth SP encounter (Face/Hand, Vices and Advice) 62.44 % from the eleventh SP encounter (Acute Angina/ Perpiheral Neuro/ Diabetes), 61.81 % from the twelfth SP encounter, and 65.24 % from the thirteenth SP encounter. As noted, there was one exception to this pattern, occurring in the analysis of the seventh SP encounter (Cough/ Non-compliance). The results for this patient encounter show that two underlying factors were extracted. However, the results of this two-factor solution were not consistent with the hypothesized relationship. That is, the items that loaded on these two factors did not correspond with the factor structure that had been hypothesized by the original authors. The first factor found in the seventh SP encounter, consisting of the items of confidence, body language, elicit information, actively listen, and careful examination, accounted for 37.27 % of the cumulative variance. A second factor, consisting of the items of rapport, empathy, and provide information, composed 30.73 % of the variance (68.00 % of the cumulative variance). Thus, overall, Hypothesis 1 was not supported.

The factor loadings for each analysis and the percentage of variance accounted for are reported in Table 3. These results reveal that the PARS is comprised of one overall factor and that the internal consistency and reliability of this factor is sound, ranging from .84 to .94 across the measurement times. These are acceptable levels of internal consistency and reliability and support the overall finding of one and not two factors when using the PARS to evaluate the interpersonal communication skills of medical students.

In order to name the factor, an examination of the particular items that loaded highest was conducted. Investigators renamed the factor the "quality of physician-patient interaction." Because the PARS is also utilized to measure other healthcare professionals' interpersonal skills and communication, the author of the current study decided to name the factor "quality of provider-patient communication." The high coefficient alpha levels support the interpretation of a total score that measures a singular construct and provides justification for interpreting one overall score. Those scoring low on this factor manifest a low quality of interpersonal communication; those scoring high demonstrate a high quality of interpersonal communication.

Table 3

*Factor Loadings for Eight Items of the PARS for 13 SP Encounters (*N = *205)*

Encounter 1: Headache	Factor loadings on single factor
Rapport 1	.886
Empathy 1	.775
Confidence 1	.859
Body language 1	.841
Elicit information 1	.824
Actively listen 1	.823
Provide information 1	.849
Careful examination 1	.776
Encounter 2: Eye Exam	**Factor loadings on single factor**
Rapport 2	.819
Empathy 2	.849
Confidence 2	.861
Body language 2	.865
Elicit information 2	.854
Actively listen 2	.873
Provide information 2	.780
Careful examination 2	.807
Encounter 3: Earache/ Sore throat	**Factor loadings on single factor**
Rapport 3	.873
Empathy 3	.825
Confidence 3	.876
Body language 3	.883
Elicit information 3	.834
Actively listen 3	.880
Provide information 3	.828
Careful examination 3	.829

Encounter 4: Chest Pain/ Shortness of Breath	Factor loadings on single factor
Rapport 4	.278
Empathy 4	.836
Confidence 4	.841
Body language 4	.847
Elicit information 4	.840
Actively listen 4	.873
Provide information 4	.855
Careful examination 4	.806

Encounter 5: Abdominal Pain/ Nausea	Factor loadings on single factor
Rapport 5	.777
Empathy 5	.787
Confidence 5	.839
Body language 5	.824
Elicit information 5	.801
Actively listen 5	.761
Provide information 5	.729
Careful examination 5	.755

Encounter 6: (OSCE) Diverticulitis/ GERD	Factor loadings on single factor
Rapport 6	.747
Empathy 6	.760
Confidence 6	.745
Body language 6	.792
Elicit information 6	.804
Actively listen 6	.787
Provide information 6	.726
Careful examination 6	.741

Encounter 7: Cough/ Non-compliance	Factor loadings on first factor of two factors	Factor loadings on second factor of two factors
Rapport 7	.310	.782
Empathy 7	.156	.883
Confidence 7	.709	.394
Body language 7	.824	.190
Elicit information 7	.697	.349
Actively listen 7	.626	.452
Provide information 7	.313	.729
Careful examination 7	.839	.134

Encounter 8: Angina / Abo. White Coat	Factor loadings on single factor
Rapport 8	.839
Empathy 8	.820
Confidence 8	.842
Body language 8	.790
Elicit information 8	.800
Actively listen 8	.823
Provide information 8	.669
Careful examination 8	.797
Encounter 9: Hip/Knee Pain Counseling	**Factor loadings on single factor**
Rapport 9	.868
Empathy 9	.833
Confidence 9	.811
Body language 9	.813
Elicit information 9	.841
Actively listen 9	.872
Provide information 9	.696
Careful examination 9	.675
Encounter 10: Face/hand, Vices and Advice	**Factor loadings on single factor**
Rapport 10	.78
Empathy 10	.74
Confidence 10	.84
Body language 10	.69
Elicit information 10	.79
Actively listen 10	.71
Provide information 10	.69
Careful examination 10	.80
Encounter 11: (OSCE) Acute Angina/ Peripheral Neuro/Diabetes	**Factor loadings on single factor**
Rapport 11	.806
Empathy 11	.810
Confidence 11	.813
Body language 11	.790
Elicit information 11	.763
Actively listen 11	.813
Provide information 11	.764
Careful examination 11	.760

Encounter 12: three 14-minute encounters based on COMLEX-USA-Level 2-PE	Factor loadings on single factor
Rapport 12	.843
Empathy 12	.751
Confidence 12	.776
Body language 12	.838
Elicit information 12	.804
Actively listen 12	.845
Provide information 12	.683
Careful examination 12	.734
Encounter 13: Third Year OSCE	Factor loadings on single factor
Rapport 13	.838
Empathy 13	.827
Confidence 13	.856
Body language 13	.825
Elicit information 13	.756
Actively listen 13	.860
Provide information 13	.768
Careful examination 13	.719

Reliability of the PARS

Hypotheses 2a and 2b, respectively, predicted that the entire PARS, as well as the two subscales, would demonstrate acceptable levels of reliability (a coefficient alpha of at least .70). Given the results of the factor analyses that were conducted to test Hypothesis 1, a reliability analysis was not performed on the two PARS subscales, but rather, the single reliability analysis included all eight items.

A separate reliability analysis was conducted for each of the 13 separate SP encounters. Across the 13 encounters, the findings revealed a coefficient alpha ranging from .84 to.95, well above the predicted .70 level (see Table 4). The obtained coefficient alphas reveal a high level of internal consistency among the items.

Table 4

Reliability of the PARS Scores across Each of the 13 SP Encounters

SP encounter (all N=205)	Cronbach's Alpha
Encounter 1	.93
Encounter 2	.94
Encounter 3	.95
Encounter 4	.84
Encounter 5	.91
Encounter 6	.89
Encounter 7	.88
Encounter 8	.92
Encounter 9	.92
Encounter 10	.89
Encounter 11	.91
Encounter 12	.91
Encounter 13	.92

Note. PARS = Professionalism Assessment Rating Scale; SP = Standardized Patient.

Quality of Provider-Patient Interaction by Class Year

The third hypothesis predicted that the mean PARS scores, now termed "quality of provider-patient communication," of DO students would increase between the first and second years and between the second and third years. To test Hypothesis 3, a one-way repeated measures ANOVA was conducted, with time (Year 1, Year 2, Year 3) as the independent variable and the overall mean PARS score (for each of the three respective years) serving as the dependent variable. The independent variable was created by averaging the performance of each student across each year, That is, the average of each student's first six encounters for Year 1 (M = 6.30), the average of the next five encounters for Year 2 (M = 6.45), and the average of the encounters that made up the third year OSCE for Year 3 (M = 6.48), as shown in Table 5. Results of the repeated measures ANOVA violated Mauchley's Test of Sphericity (W = .87, Chi-square = 28.69), but the Greenhouse-Geisser correction was utilized and indicated significant differences over time (Greenhouse-Geisser = .88, df = 2), as shown in Table 6.

Table 5

Mean PARS Scores by Class Year

Class Year	Mean	Standard Deviation
1	6.30	.45
2	6.45	.39
3	6.48	.58

Results of the analysis indicated that performance improved across each of the 3 years. A post hoc analysis of the comparisons of the previously stated means showed the students' performances at Year 2 and Year 3 were significantly higher than their performance at Year 1. However, there was no difference between Years 2 and 3. Clearly, there was an improvement in the performance over the 3 years to show significant improvement over time.

Table 6

ANOVA Summary Table

Source		Sum of Squares	df	Mean Square	F
Factor 1	Greenhouse-Geisser	3.36	1.84	1.82	11.44***
Error (factor 1)	Greenhouse-Geisser	59.91	376.19	.16	
Total		63.27	378.03	1.98	

***$p < 0.0001$

Correlation by Class Year

Hypothesis 4 proposed that there would be a significant positive correlation between first-, second-, and third-year PARS scores and that the relationship would be linear. To test this hypothesis, a Pearson product moment correlation was conducted between mean PARS scores for Years 1, 2 and 3 to identify any linear relationship. All three of these relationships were both positive and significant. That is, there was a

significant positive correlation between the performance of the students on the PARS between years one and two, between Years 1 and 3, and as well between Years 2 and 3. Coefficients of determination were calculated for each of the reported correlations. These data reveal that for the correlation of .47 between Year 1 and Year 2, the coefficient of determination was .22, meaning that 22 % of the variability in performance at Year 2 is attributable to differences in performance at Year 1. The remaining 78 % of the variability is unaccounted for and is the result of any number of factors.

The correlation between performance at Year 2 and 3 was .42; the associated coefficient of determination was .1789, meaning that 17.89 % of the variability in performance in Year 3 is attributable to differences in performance during Year 2. The remaining 78.11 % of the variability is unaccounted for and is the result of unknown factors.

The correlation between performance at year 1 and 3 was .36; the associated coefficient of determination was .1289, meaning that 12.89% of the variability in performance in Year 3 is attributable to differences in performance during Year 1. The remaining 87.11% of the variability in performance at Year 3 is left unaccounted for and also the result of unknown factors.

Table 7

Correlation Matrix across Years of Medical School

	First Year	Second Year	Third Year
First Year		.470**	.359**
Second Year			.423**

**$p < 0.001$

Relationship between PARS Scores and Ratings by Clerkship Supervisors

Hypothesis 5 indicated that there would be a positive correlation between DO students' mean (overall) PARS scores and the mean of the DO students' clerkship supervisors' ratings. This hypothesis was tested by examining the relationship between average PARS scores for Years 1, 2 and 3 and the clerkship ratings for those same years. The results are shown in Table 8. Clerkship ratings were significantly correlated with performance on the PARS on only three of the selected correlations studied. The remaining correlations are reported for consideration but should be viewed with caution as these results may capitalize on chance.

Table 8

Pearson Correlation Matrix between Mean PARS scores by Year and Selected Clerkship Ratings

	Family Medicine	Pediatrics	OBGYN
PARS Year 2	.154**	.173*	
PARS Year 3			.153**

*$p < .01$, **$p <$. .015

Pearson product moment correlations were calculated between performance on PARS and Clerkship ratings in selected rotations. Significant positive correlations were noted between Year 2 PARS ratings and the family medicine clerkship ratings, $r = .154$, $p < .014$, $N = 204$, between Year 2 PARS ratings and the pediatric clerkship ratings, $r = .173$, $p < .007$, $N = 204$, and Year 3 PARS ratings and the OBGYN rotation $r = .153$, $p < .015$, $N = 204$. The higher the PARS rating, the higher the clerkship rating, and the lower the PARS rating, the lower the clerkship rating for these years and rotations. A coefficient of determination denotes the strength of the linear relationship between two variables. This statistic related to the strength of the relationship between the preceptor rating for the family medicine clerkship and the Year 2 PARS rating was 2.37. This statistic indicates that 2.37 % of the preceptor rating for the family medicine clerkship was attributable to differences in PARS ratings during the second year. The coefficient of determination for the relationship between the preceptor rating for the pediatrics clerkship and the Year 2 PARS rating was 2.99. This statistic indicates that 2.99 % of the variability in the preceptor rating for the pediatric clerkship was attributable to differences in PARS ratings during the second year. The coefficient of determination for the relationship between the preceptor rating for the OBGYN clerkship and the Year 2 PARS ratings was 2.34. This statistic reveals that 2.34 % of the variability in the preceptor rating for the OBGYN clerkship was attributable to differences in PARS scores during the second year. Therefore, while clerkship ratings were found to be correlated in some cases to the PARS scores, they consistently accounted for a small portion of the variance.

Chapter 5: Discussion

Implications of Results

The present study examined the psychometric properties of the Professionalism Assessment Rating Scale (PARS) in relation to its utilization as a measure of the quality with which Doctor of osteopathy (DO) students completed 13 standardized patient (SP) encounters over the course of their t3 years of medical school. I will first discuss the implications of the results for each of the five hypotheses, specifically (a) the single factor structure of the PARS, (b) the coefficient alpha reliability of the PARS, (c) the increase in PARS scores between first and second years and between second and third years, (d) the positive correlations between first-, second-, and third-year PARS scores, and, finally (e) the correlation between the DO students' PARS scores and the DO students' clerkship supervisors' ratings for the students' first, second, and third years of osteopathic medical school. Next, the limitations of this study and ways in which these limitation may be remedied will be discussed. Clinical and educational implications of the findings will then be discussed. Finally, future research will be suggested.

Quality as Single Factor Measured by the PARS

Regarding Hypothesis 1, the results of the factor analyses, as previously noted, indicate that the PARS is comprised of a single factor herein named "the quality of provider-patient interaction." In retrospect, the fact that two distinct factors did not emerge is not extremely surprising, nor is a two-factor solution absolutely necessary in order to validate the PARS as a reasonable measure of medical students' interpersonal interaction skills and learning.

As originally designed, the PARS was created to measure two different dimensions; however, empirical data do not support a two-factor solution. Although the factor loadings across all but one SP were uniformly high, the most parsimonious explanation for the results is probably that there are simply not two factors to the PARS. The PARS, in overall terms, would appear to be appraising provider-patient interaction in a rather holistic fashion. Overall quality of provider-patient interaction is then perhaps a singular factor, and this single factor is best illuminated by all eight items taken as a whole. Thus, while not supporting the hypothesis, this result does not undermine the overall purpose of this study – that of exploring and validating the psychometric properties of the PARS. Based on the view of the original test developers of the PARS, the expectation was that there would be two different dimensions to the PARS. This expectation turned out not to be the case.

Based on the empirical findings herein, one might simply conjecture that the quality with which the DO students interact with the SPs is relatively consistent across all eight rated items of the PARS. Perhaps, from the perspective of the SP, the quality of the examination provided to the patient (the quality of the questioning, the quality of the listening, the quality of the counseling, and the quality by which a thorough, careful physical examination or treatment is provided) and the quality of the relationship established with the patient (establishment and maintenance of rapport, demonstration of empathy and confidence, and the use of appropriate body language) cannot be separated, but rather work together in concert to create a holistically felt sense of quality in the interaction. Indeed, this idea is supported by the literature on the provider-patient

working alliance, which suggests that the quality with which a provider-patient interaction occurs is, in and of itself, a skill that parallels and is ultimately interwoven with technical and knowledge-based skills. This idea is discussed later as it relates to the concept of *habits of mind* (also referred to as *mindfulness*).

The emergence of the singular factor, the quality of provider-patient interaction, across all but one encounter also speaks to the robustness of the PARS as an instrument that can be used in measuring this quality of interaction in both summative and formative capacities. For example, while the results herein did not support a two-factor solution, it one should note that the results consistently and clearly supported a unitary factor. For instance, suppose that the factor analysis results has been highly inconsistent, sometimes supporting a unitary factor, and sometimes supporting diverse underlying factors. How might this finding be interpreted relative to the important goal of legitimating PARS as a valid measure of DO student skill? Such invariance across measurement periods supports the consistency of the factor being measured by the PARS. Moreover, the factor analytic results are further supported by the results of the reliability analysis.

Reliability of the Unitary Concept Measured by the PARS

Hypothesis 2 proposed that the PARS would prove to be a reliable instrument. Examination of coefficient alphas for the PARS across all 13 encounters shows them to be uniformly high and acceptable. These results strongly support the existence of a unitary construct as opposed to the proposed two-factor PARS. The PARS thus appears to be a highly reliable measure of the overall quality of provider-patient interaction. High coefficient alphas support internal consistency and strongly support the idea that the

PARS measures a singular concept. This finding also supports the use of one overall score. Many of the explanations, definitions, and descriptions of the physician working alliance allude to interwoven and overlapping constructs, such as interpersonal skills, professionalism, ethical principles, communication skills, and cultural constructs (Boulet et al., 1998; Ong et al., 1995; Van Zanten et al., 2005; Weidner, Gimpel, Boulet, & Solomon, 2010). It is thus not surprising that the PARS, as a unitary factor, is shown to be highly reliable in measuring the quality of provider-patient interaction – an element so integral to the working alliance.

In sum, the single-factor structure found for the PARS perhaps reflects the difficult task of empirical measurement relative to the quality of any humanistic skill. Relative to the hypothesized two-factor construct for the PARS items, the separation of the scale into two factors is unwarranted, and there is no empirical basis for separating the items into two factors. For the purposes of outcome measurement, the overall scores on the PARS may be quite useful as a measure of interpersonal skills. Feedback about performance on the eight items may prove useful talking points that the SPs can include in feedback on the quality of the DO students' provider-patient interaction skills.

Development of DO Student Skill Over Time as Measured by the PARS

Hypothesis 3 proposed that DO students' PARS scores would improve over time. Given the ANOVA results, it can be said with confidence that the quality of the DO students' provider-patient interaction does improve over time. These results clearly lend support to the idea of the PARS as an instrument for measuring the learning DO students experience at PCOM. In fact, the DO students' improvement from Year 1 to Year 2 and

Year 1 to Year 3 is consistent with a developmental model of medical education and learning. That is, these improvements in performance over time suggest that significant learning and acquisition of skills do transpire, relative to the quality with which DO students at PCOM interact with the SPs during their regularly scheduled SP encounters. This theory is consistent with the model of professional competence put forth by Epstein and Hundert (2002).

The strong support for the Hypothesis 3, which revealed an increase in professional interaction, as measured by the PARS, from the first to second and the first to third year of DO training is important. The significant difference from the first to second year and the first to third year suggests that students at PCOM do, in fact, across time develop in their ability to build a working alliance with their patients, while essentially also learning increasingly difficult technical skills and maintaining highly rigorous schedules.

Consistency in Aptitude as Measured by the PARS

Hypothesis 4 shows PARS scores across Year 1 through Year 3 to be positively and significantly correlated with one another. The fact that the scores of the student physicians were positively correlated across years indicates a consistent pattern of performance on the standardized exams across time. That is, for physicians in training who did well (or poorly) in one year, their scores in ensuing years were to some significant degree consistent. This result, along with the results of the reliability analyses, appears to indicate that the PARS is a useful instrument for measuring the quality of DO students' provider-patient interactions in SP encounters. The significant

positive correlations between the first, second, and third years also speak well of PCOM's training of DO students. This positive correlation between years of training suggests consistency in performance across years.

The amount of unaccounted variance evidenced by the obtained coefficients of determination may be attributable to any of a number of phenomena. Other possible factors associated with scores may be related to individual differences in characteristics in DO students or in SP ratings. Moreover, the increasing amount of technical skill necessary to successfully perform the treatments required in the SP examinations may divert attention away from interpersonal awareness, particularly during an encounter that requires an especially difficult technical procedure. Further, the technical skill required for the encounters increases over time, mirroring the increasing difficulty of the course material as the DO students' progress through medical school. Elements of medical-student stress and burnout have also been named as having an effect on medical student professionalism and empathy (Bellini & Shea, 2005; Diseker & Michielutte, 1981; Hojat et al., 2004; Shanafelt et al., 2005). Despite the unaccounted for variance, this study did show the DO students to consistently improve in the quality with which they communicated with their patients over time. The identification of the moderators of that improvement and learning represents a valuable inquiry to be made using PARS scores and SP encounters in the future.

Construct Validation of the PARS

The investigation of the relationship between PARS scores and Clinical Competency Evaluations (CCE; clerkship supervisor ratings) proved significant in a few

cases as previously reported. These significant correlations suggest the PARS is to some limited degree a valid measure and can be assumed to measure that which it claims to measure—the quality of the provider-patient interaction. The significance in these correlations may result more so from three of the seven osteopathic competencies that most relate to the items of the PARS, as described in Chapter 3 and discussed later in the "Limitations" section. Factors other than professionalism (e.g., achievement test score on clerkship content) contribute to the scores on the clerkship ratings and as such may have attenuated the correlations. Thus, the small account of the overall variance may be due to the scope of the measures differing, a limitation of this study discussed below. Future researchers may choose to compare the PARS scores to clerkship supervisors' ratings pertaining to specific competencies more closely related to the items in the PARS. To be specific, clerkship supervisor evaluations address all of the osteopathic competencies, whereas the PARS most closely maps to only three. That the correlations of PARS scores to clerkship supervisor ratings were, in overall times, quite weak is not surprising; the coefficients of determination here suggest a large amount of unaccounted for variance.

Limitations

Further examination of the PARS, as well as of other potential measures of construct validity, should be pursued in future research. This section will discuss limitations of the current study, including quality assurance of SPs, the prior knowledge of elements in the PARS, the differences between SP and real patient encounters in the evaluation of behavior in provider-patient encounters, and the time duration of the study.

Use of CCE for Construct Validation of the PARS

A comparison between the PARS scores and CCE ratings in pursuit of construct validation of the PARS warrants further investigation. First, one must note that the reliability of the clerkship ratings has not been examined. That is, no empirical evidence exists to legitimate the reliability of the ratings across clerkship supervisors. Simply put, a number of factors were not sufficiently controlled for in regard to the clerkship ratings. While "the quality of provider-patient interaction" may be implied in the clerkship evaluations, in actuality the clerkship evaluations are likely more closely related to clerkship supervisors' ratings of the specific osteopathic competencies of Patient Care, Professionalism, and Interpersonal and Interaction Skills. Also, the clerkship ratings do not specifically target the quality with which skills are performed, nor are they completed from the perspective of the patient. Alternatively, the low rate of correlation may be caused by the nature of the clerkship ratings themselves, in that they were highly dependent on a written examination and only partly determined by actual observation of the practitioner-patient interpersonal interaction, let alone the quality of that interaction.

Furthermore, the faculty members who observed physician encounters may or may not have looked for the quality of the interaction. In addition, the attention paid to quality may well have depended upon the types of procedures and patients typical of each rotation. Finally, because the faculty members observed physician encounters from outside of the provider-patient interaction, they did not intrinsically experience the same perspective as that of the SPs who awarded the PARS scores. Accordingly, the large

amount of unaccounted for variance between the DO students' PARS scores and the ratings given by their clerkship supervisors cannot necessarily be taken to indicate that there is no relation between DO students PARS scores and the quality of performance in their clerkships.

Quality Assurance of the SPs

Caution should be taken when using this type of measure because it is subject to the biases and preferences of those conducting the assessment, in the present case the SPs who completed the rating scale. While quality assurance efforts were undertaken during the training of the SPs, relative to their rating of the students on the PARS, the SPs' responses were not explicitly monitored or documented during the SP encounters at PCOM. The National Board of Osteopathic Medical Examiners (NBOME) uses SPs in the administration of the Comprehensive Osteopathic Medical Licensing Examination (COMLEX) Performance Evaluation (PE), which is used to assess the competence of the DO candidates for licensure. Candidates must exhibit professionalism in interpersonal and interactional skills, as well as in technical skills, in order to successfully complete these examinations. Proper training of SPs, along with repeated simulated encounters, has shown that scoring can be reproduced and provides evidence to suggest that performance with SPs can be translated to live clinical settings (Tamblyn et al., 2007). Further, rater bias can be minimized through careful recruitment and training of SPs. Mathematical adjustments to individual raters' stringency or leniency also can be performed, so as to standardize these ratings. An example of this type of quality assurance is described in the preparation of the SPs for the COMLEX-USA Level 2–PE.

In preparation for rating DO candidates for licensure in the COMLEX-USA Level 2-PE, SPs are given a comprehensive training for quality assurance purposes (Wiedner et al., 2013). First, the SPs are provided an overview of the osteopathic competencies and the six domains of the Global Patient Assessment (GPA) shown in Table 9. The GPA is the global rating scale that is used to measure the DO candidates' humanistic skills in the licensing examination. Next, the SPs are shown examples of observable behaviors that represent "unacceptable" and "superior" ratings for each of the GPA's six domains. They then move on to identify these observable behaviors in videos in terms of the six GPA domains. Subsequently, the SPs discuss the rationales for their given ratings and a benchmark score is established for each dimension in the training videos. The SPs must score within 1 point of the benchmark score before moving to the next phase of the training. The next phase consists of the SPs participating in a pretesting station in an actual exam with a candidate. SPs who are approved to participate in the exam are monitored for consistency and accuracy in their rating of the candidates. The scoring and performance of 10% of each of the SP's encounters are reviewed for quality assurance (Weidner et al., 2010). Daily analyses of means, standard deviations, frequencies, and correlations among GPA ratings are completed, and monthly aggregate data are reviewed for discrepancies and changes in scoring patterns. SPs who vary too far from the norm in their scoring are required to participate in remediation training before they can resume their roles. Candidates rated below acceptable levels are reviewed by doctoral-level staff. Refresher trainings are required each year for all SPs' assurance (Weidner et al., 2010).

Table 9

The Six Domains of the Global Patient Assessment (GPA)

Elicit information	Ability to ask questions in an articulate, understandable, straightforward manner; Appropriate use of open-ended questions and facilitative prompts A good command of spoken English
Listening skills	Ability to both listen and respond appropriately to patients statements and questions
Giving information	Ability to effectively communicate clear explanations and counseling with regard to the patients' concerns and to facilitate appropriate closure.
Respectfulness	Ability to treat the patient in a polite and sensitive manner, honoring other's choices and rights, demonstrating appropriate cultural competence, holding the patient in high regards or esteem.
Empathy	Ability to demonstrate and communicate (verbally /nonverbally) understanding, concern and interest in the patient’s medical problem and life situation.
Professionalism	Ability to appear both appropriately confident and therapeutic, to show altruistic interest in the patient's welfare and to ensure patient confidentiality and high ethical principles (American Osteopathic Association, Core Competency…, n.d.).

Effects of Prior Knowledge of Elements of Evaluation on Performance

This point is a rather obvious but nevertheless important. Students were aware of being evaluated. This awareness could in and of itself affect their performance on the SP exams. Specifically, this observer effect would be expected to enhance the students' performance and resultant PARS scores. This limitation also applies to the clerkship ratings, insofar as student physicians knew that they were being evaluated by their supervisors relative to their performance.

Differences between SPs and Real Patients in Simulations

SPs have long been utilized in the evaluation of medical students and provide a good control for the variance in presentation that accompanies real patients (Stimmel, Cohen, Fallar, & Smith, 2006). Faculty members also have long been used as evaluators of medical student-patient interactions. The observation effect (reactivity) explains that individuals will perform differently when they know they are being observed. Still, past research evaluating medical students' performance after communication skills training shows similar behavior with real patients and SPs and improvement with both as well (Delvaux et al., 2004; Imel et al., 2014), even though ratings by real patients are generally lower than those by SPs (Reinders et al., 2011). It is possible that SP evaluations are more objective (Cleland, Abe, & Rethans, 2009, Fiscella, Franks, Srinivasan, Kravitz, & Epstein, 2007), and the SP's knowledge of the encounter prior to its occurrence is thought by some researchers to be an advantage in that it allows for more pertinent feedback to students (Bokken et al., 2010). That being said, SP encounters are simulations and therefore inherently different from real patient real patient encounters (Cleland et al.,

2009). Recent research by Oliver, Hausdorf, Lievens, and Conlon (2014) sheds light on the moderators of such differences. Their study looked at 184 interpersonal simulations and found that differences in the portrayed disposition of the SPs effected the relationship building and directive communication behavior of the individuals being evaluated. In addition, the same differences in portrayed disposition moderated between the use of these interpersonal behaviors and performance ratings. Thus, differences in individual patients and individual medical students will inevitably affect resulting scores on such interpersonal measures as the PARS. The use of SPs allows for greater control in order to compare participants to each other and across time, and in order to train medical students in specific skills (Imel et al., 2014).

Time Duration of the Study

One final limitation that merits mention deals with the relatively limited time duration of the study. The student physicians in this case were tracked during their years of osteopathic training, as well as during their subsequent clerkships. However, in reality, the careers of medical professionals span a number of decades. The question of how physicians perform in the long run along the parameters used in the study is important. For example, researchers propose that medical students' ability to create a strong working alliance is less well developed than more seasoned physicians and, therefore, more susceptible to stress levels (Borrell-Carrio' & Epstein, 2004). Poor professional behavior in medical school has been shown to be a predictor of censure by licensing boards (Moffat et al., 2004; Papadakis et al., 2005). Perhaps the quality of interaction, when well developed and habitual, can provide a buffer to stress or burnout

that might otherwise occur later in a physician's career. As such, the formative use of the PARS in medical school may positively affect the degree to which physicians trained to attend to the quality of their provider-patient interactions are able to affect a positive physician-patient alliance. Or, perhaps students who are sensitized during their medical education to the importance of the practices of empathy, active listening, providing a careful physical exams, and so on develop a positive skill set and mindset that follow them and are well utilized throughout the length of their careers. These points are all speculative but are important questions to pose, and as such, they form part of the limitations of this study.

Educational Implications of the PARS

Having found the PARS to be comprised of a single and highly internally consistent factor, the PARS appears to be a valid and reliable measure that can usefully be employed for both formative (developmental/educational) and summative (evaluative) ends. For example, the creators of the GPA suggest that it comprises a high-stakes summative examination, designed to measure minimal levels of competence, and as such, it is not to be used for teaching, formative assessment, or best practices (Weidner et al., 2010). In contrast, the creators of the PARS designed the PARS specifically for use as a formative assessment, with the expectation that the SPs would give informed feedback to DO students on many of the SP encounters in which they participate.

Currently, once each year, the SP encounters are used as a summative measure. The author of this study suggests that the strong reliability of this measure supports its wider use as global assessment scale for summative purposes more generally. Thus, the

suggestion that the PARS continue to be used, like the GPA, as a summative global assessment. At the same time, however, we also recommend that individual items be kept as teaching points when the PARS is used at PCOM, and in this way it can also be used in a formative capacity. Thus, in overall terms, the PARS can be used for both formative and summative purposes. An increase or decrease in PARS scores may be indicative of the quality of curriculum, role modeling, personal development of the DO students, and developmental level of competence.

Professional Competence as Measured by the PARS

Professional competence is composed of technical skill, knowledge, and moral development (Epstein & Hundert, 2002). It can be conceptualized as including (a) cognitive functions such as the use and acquisition of knowledge to conceptualize a patient's case; (b) integrative functions, such as those necessary to incorporate psychosocial and biomedical data in clinical reasoning; (c) relational functions, such as effective interaction with patients, community resources, and with colleagues; and (d) affective/moral functions, such as the personal awareness to use these skills with patience and with clear judgment. "Professional competence is developmental, impermanent, and context-dependent," claimed Epstein and Hundert (2002). Thus, one might argue that healthcare providers who make a habit of being present with and attentive to their patients are able to devote a greater portion of their mental focus to the technical skills and clinical reasoning that are necessary to accurately diagnose and treat patients. Students who possess such *habits of mind* also likely attain and realize high levels of quality in interactions with their patients. These students may therefore maintain a strong

working alliance despite the level of focus they may need to place on a technical procedure that is new to them.

Development of Professional Competence Using the PARS

A developmental model describes professionalism on a continuum ranging from that of the individual to that of a complex system of interpersonal relationships (Batalden et al., 2002; Hilton & Slotnick, 2005). This model is relevant to the training of physicians because it suggests that professional interaction skills can be developed over time. Efforts to assess medical professionalism have focused on technical skill and tend to overlook other key components, such as the integration of knowledge and practice, lifelong learning, and interpersonal skills (Epstein & Hundert, 2002).

The science of evidence-based medicine is rooted in the acquisition and generation of knowledge, problem solving and decision-making in the context of the clinical setting (Epstein & Hundert, 2002). As physicians' development unfolds, what was once explicit knowledge becomes integrated with the physicians' experiences to become what is known as tacit knowledge. Tacit knowledge includes the intentional use of heuristics, intuition, and pattern recognition (Polanyi, 1969). As medical students develop into experts, clinical skills, including the use of heuristics, are integrated into the physicians' personal knowledge and become increasingly difficult to identify in a formal assessment. Personal knowledge does not always lead to right judgment, and therefore physicians must develop habits of mind that include emotional as well as cognitive self-awareness in order to monitor their own biases.

Distinct competencies work together, combining to create the whole of professional competence. Expertise in one specialty of medicine, for instance, does not translate to expertise in another. Expertise is more than knowledge of treatment of biological systems; it requires physicians to be confident enough to tolerate uncertainty and to make decisions in ambiguous situations (Schon, 1983). The support that this study found for Hypotheses 3 and 4 may indicate that PCOM students are developing habits of mind as they progress through their training. Hypothesis 3 suggests that the increase in technical skill necessary for more difficult tasks did not interfere with the quality of the students' provider-patient interactions. Hypothesis 4 suggests that PCOM training provides the support for all DO students to improve in their ability to form a working alliance over time.

Clinical Implications of the PARS

The development of instruments to measure aspects of the quality of provider-patient interaction (interpersonal skills and interaction) has become an integral component of evaluations of professional competence, broadly speaking. As discussed previously, research shows the importance of the working alliance to patient outcomes, to patient adherence, and to patient and provider satisfaction, as well as to malpractice concerns (Adamson et al., 1989). These research findings, coupled with the changing nature of the healthcare system in the United States, make it vital that physicians enjoy high-quality interactions with their patients, quality interactions that must transpire within the often quite limited time they have available.

For example, research on empathy alone (a component of the working alliance and an item on the PARS) shows high levels of communication and interpersonal skill to be correlated with empathy (Langenau et al., 2011). Research further shows that these communication and interpersonal skills improve when individuals are specifically trained in empathy skills (Feighny et al., 1995), and also shows that empathy is directly linked to positive patient outcomes (Hojat, Louis, Markham… et al., 2011; Moore, Wilkinson, Rivera, & Mercado, 2004). After completing a communication skills training, individuals' self-ratings of clinical competence improve, as do observations of their confidence by SPs and faculty member observers (Carvalho et al., 2014; Delvaux et al., 2004). Likewise, patients report a higher level of perceived sense of care from their physicians with training in communication skills, (Moore et al., 2013). Also, a distinct difference between the simple behavior of performing a specific skill and the quality with which such a given skill is applied is important. The quality with which a particular skill is applied can make that skill fundamentally different from the same skill in its rudimentary form. Moreover, the more qualitative facets of certain skills, the more vexing those skills are to describe and to teach to others in an objective manner. Because the teaching and assessment of such qualitative elements can be more difficult and ambiguous, these elements are often left out of the curriculum. The PARS provides a measure that not only supports the pedagogy of elements that are often more qualitative in nature, but also provides a means for assessment of students' acquisition of these important skills.

It is also important to note that models indicate that medical students' performance in medical school is predictive of their subsequent professional behavior in both residency and their future practice (Hojat et al., 2009). In this way, performance during medical training is realistically and explicitly linked to patient outcomes once students graduate and engage in practice within the changing healthcare system.

The emerging changes in the field of health psychology and the growth of the patient-centered medical home model require the interdependence of healthcare providers across disciplines. The healthcare system, in its increasing movement toward bundled payments, requires the integration of behavioral health components. Thus, the more informed each provider is of the multiple factors involved in each patient's condition, the more likely the providers will be able to appropriately diagnose and treat each patient. In other words, the sharing of information among professionals can prevent time spent in treatments for diagnoses that are not informed by the complete picture. Doctoral-level psychologists trained to work with medical professionals and educated to understand not only psychological but also basic physiological, neurological, social, and behavioral contributions to disease and healing can utilize their understanding to inform, train, and collaborate with physicians and other medical staff.

Future Research

The aforementioned benefits of the PARS scale in the education of medical students and healthcare professionals are numerous. Establishing the construct validation of the PARS will be important for medical educators to continue to use the scale with confidence. Future research might also establish the PARS as a predictive measure of

medical students' professional behaviors and career paths. Students in other healthcare fields may benefit from the use of the PARS for evaluation and feedback. Finally, the investigation into the quality of provider-patient interaction, as measured by the PARS, may be used to investigate quality as a moderator of a strong working alliance, patient outcomes, adherence, and malpractice.

Further Construct Validation of the PARS

In pursuit of construct validation of the PARS, other comparable measures may be sought for comparison. Researchers may, as previously discussed, wish to dissect the clerkship ratings in order to examine only those specific competencies most closely related to the skills that the PARS deems to measure. Alternatively, researchers may seek other measures, such as the GPA used in the COMLEX-USA Level 2-PE to compare with the PARS for the purpose of construct validation.

Researchers may also seek validation of the PARS by breaking out the elements of the clerkship ratings so as to compare the osteopathic competencies most closely related to the quality of provider-patient interaction. The seven AOA professional competencies include three that are particularly relevant to those skills found in the eight items of the PARS.

Another measure that may be used for future efforts to validate the PARS could be DO students' performance on the GPA previously discussed as it is used in the COMLEX-USA Level 2-PE. The GPA is a global rating scale like the PARS, which is likewise used to rate student performance (Weidner et al., 2010). It uses operational definitions to achieve a holistic measure of DO candidates' humanistic skills. Zhang and

Roberts (2103) have shown it to be a reliable instrument when used in a simulated environment with SPs. The GPA is much like the PARS in that a 9-point Likert-type scale is used by SPs trained to rate the candidates based on the sum total of relevant behaviors during each encounter. On this scale an *unacceptable performance* is rated 1-3, an *adequate performance* is rated 4-6, and a *superior performance* is rated 7-9. The six domains of the GPA and their operational definitions are shown in Table 9.

PARS as a Predictive Measure

Research discussed previously links medical students' professionalism and elements thereof to future behavior in medical school and beyond. Elements of the PARS, such as empathy, have been established as predictors of future behavior, although inconsistently (decreases in clinical years, allopathic school, Chen, Lew, Hershman, & Orlander, 2007; remained constant, osteopathic school, Hojat et al., 2002; Kimmelman, et al., 2012; supervisor ratings remain constant, Colliver, Conlee, Verhulst, & Dorsey, 2010); higher in third year than allopathic counterparts, Calabrese, Bianco, Mann, Massello, & Hojat, 2013).

PARS Used with Broader Range of Healthcare Providers

The PARS as is described and reported herein is used in SP encounters with DO students. It can as easily be applied to such encounters in the training of other healthcare professionals such as nursing students, physician assistants, and psychology students.

Investigation of Quality as Moderator of Other Behaviors

As previously stated, there is a good amount of unaccounted for variance, and therefore, unaccounted for factors to which the consistent improvement of the DO

students' PARS ratings may be attributed. The identification of the moderators of that improvement and learning represents a valuable inquiry to be made in the future. Owing to the clinical implications and significance of the provider-patient relationship on patient outcomes, adherence, malpractice, and patient satisfaction, this pursuit is worthy of future research.

Summary

This study examined the psychometric properties of the Professionalism Assessment Rating Scale. Results showed the PARS to be made of a single factor for which all eight items correlated strongly. The PARS is a highly reliable construct measuring the quality of the provider-patient interaction during SP encounters over the course of the first 3 years of the osteopathic medical education at PCOM. Followed over the course of 3 year, PARS scores improved across time, suggesting the students at PCOM are, in fact, developing interpersonal and communications skills as they proceed through their education. PCOM is likely providing formative experiences by which the students improve each year, part of which are the SP encounters and feedback provided to DO students during their training. This study demonstrates the PARS to be internally reliable and construct valid evidenced by its factor structure. However, less robust is the construct validation found in this study, which is likely because the clerkship ratings, proposed to correlate with PARS scores, measure a broader range of skill than the skills the PARS encompasses.

The PARS is a robust instrument that has been utilized in SP encounters at PCOM for many years. Limitations to this study do not take away from the clinical and

educational implications named in this discussion. Limitations of this study include the need for quality assurance among the SPs who use the PARS measure to rate the quality of the DO students' provider-patient relationship, differences between SPs and real patients, and the brief time duration of the study. Implications for educational use of the PARS include both formative feedback given by SPs and faculty members and summative evaluation. The PARS provides an important measure of professional competence that can be monitored over the course of the DO students' education to ensure they are progressing in their knowledge and skill as they proceed through increasingly complex SP encounters. Specifically, the PARS touches upon the habits of mind that must be developed in medical students as they incorporate tacit knowledge with technical and interpersonal skill in their patient interactions. Numerous links exist between the quality by which physicians interact with their patients and patient satisfaction, adherence, malpractice, and patient outcomes. These skills necessary to building a strong working alliance are important in the current healthcare system, which requires physicians to gather more information in less time to diagnose and treat their patients.

The investigator in this study suggests options for future research toward the construct validation of the PARS and warrants it as an important pursuit because of the clinical and educational implication of the use of the PARS. The PARS may be used as a predictive measure in the future. It also may be used with a broader range of healthcare providers. Finally, in the future, researchers may wish to investigate the possibility of quality of interaction as a moderator for success in other behaviors. The PARS provides

a measure of the quality of provider-patient interaction that is internally consistent, and can be used in the teaching and evaluation of interpersonal skills of healthcare professions—skills so vital to patient outcomes.

References

Adamson, T. E., Tschann, J. M., Gullion, D. S., & Oppenberg, A. A. (1989). Physician communication skills and malpractice claims. A complex relationship. *Western Journal of Medicine, 150*(3), 356-360.

Ainsworth, M. A., & Szauter, K. M. (2006). Medical student professionalism: Are we measuring the right behaviors? A comparison of professional lapses by students and physicians. *Academic Medicine, 81*(10), S83-S86.

Aktekin, M., Karaman, T., Senol, Y. Y., Erdem, S., Erengin, H., & Akaydin, M. (2001). Anxiety, depression and stressful life events among medical students: A prospective study in Antalya, Turkey. *Medical Education, 35*(1), 12-17.

Alexander, J. A., Hearld, L. R., Mittler, J. N., & Harvey, J. (2012). Patient-physician role relationships and patient activation among individuals with chronic illness. *Health Services Research, 47*(3), 1201-1223. doi:10.1111/j.1475-6773.2011.01354.x

Ambuel, B., & Mazzone, M. F. (2001). Breaking bad news and discussing death. *Primary Care: Clinics in Office Practice, 28*(2), 249-266.

American Osteopathic Association. (2010). *Core Competency Compliance Program Part 1.* Retrieved January 7, 2013, from http://www.osteopathic.org/inside-aoa/accreditation/postdoctoral-training-approval/Documents/core-competency-compliance-program-part-1.pdf

American Osteopathic Association. Osteopathic Oath. Retrieved March 3, 2015 from http://www.osteopathic.org/inside-aoa/about/leadership/Pages/osteopathic-oath.aspx

Batalden, P., Leach, D., Swing, S., Dreyfus, H., & Dreyfus, S. (2002). General competencies and accreditation in graduate medical education. *Health Affairs, 21*(5), 103-111.

Beck, R. S., Daughridge, R., & Sloane, P. D. (2002). Physician-patient communication in the primary care office: A systematic review. *The Journal of the American Board of Family Practice, 15*(1), 25-38.

Becker, E. R., & Roblin, D. W. (2008). Translating primary care practice climate into patient activation: The role of patient trust in physician. *Medical Care, 46*(8), 795-805.

Bell, R. A., Kravitz, R. L., Thom, D., Krupat, E., & Azari, R. (2002). Unmet expectations for care and the patient-physician relationship. *Journal of General Internal Medicine, 17*(11), 817-824.

Bellini, L. M., & Shea, J. A. (2005). Mood change and empathy decline persist during three years of internal medicine training. *Academic Medicine, 80*(2), 164-167.

Bennett, J. K., Fuertes, J. N., Keitel, M., & Phillips, R. (2011). The role of patient attachment and working alliance on patient adherence, satisfaction, and health-related quality of life in lupus treatment. *Patient education and counseling, 85*(1), 53-59.

Berry, L. L., Parish, J. T., Janakiraman, R., Ogburn-Russell, L., Couchman, G. R., Rayburn, W. L., & Grisel, J. (2008). Patients' commitment to their primary physician and why it matters. *The Annals of Family Medicine, 6*(1), 6-13.

Blanquicett, C., Amsbary, J. H., Mills, C., & Powell, L. (2007). Examining the perceptions of doctor-patient communication. *Human Communication, 10*(4), 421-436.

Blasi, Z. D., Harkness, E., Ernst, E., Georgiou, A., & Kleijnen, J. (2001). Influence of context effects on health outcomes: A systematic review. *The Lancet*, *357*(9258), 757-762.

Bokken, L., Rethans, J. J., Jöbsis, Q., Duvivier, R., Scherpbier, A., & van der Vleuten, C. (2010). Instructiveness of real patients and simulated patients in undergraduate medical education: A randomized experiment. *Academic Medicine*, *85*(1), 148-154.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, *16*(3), 252-260. doi:10.1037/h0085885

Borrell-Carrió, F., & Epstein, R. M. (2004). Preventing errors in clinical practice: A call for self-awareness. *The Annals of Family Medicine, 2*(4), 310-316.

Boulet, J. R., Ben-David, M. F., Ziv, A., Burdick, W. P., Curtis, M., Peitzman, S., & Gary, N. E. (1998). High-stakes examinations: What do we know about measurement? Using standardized patients to assess the interpersonal skills of physicians. *Academic Medicine, 73*(10), S94-S96.

Boulet, J. R., & Swanson, D. B. (2004). Psychometric challenges of using simulations for high-stakes assessment. In W. F. Dunn (Ed.), *Simulators in Critical Care Education and Beyond* (pp. 119-130). Des Plaines, IL: Society of Critical Care Medicine.

Calabrese, L. H., Bianco, J. A., Mann, D., Massello, D., & Hojat, M. (2013). Correlates and changes in empathy and attitudes toward interprofessional collaboration in osteopathic medical students. *JAOA: Journal of the American Osteopathic Association*, *113*(12), 898-907.

Carvalho, I. P., Pais, V. G., Silva, F. R., Martins, R., Figueiredo-Braga, M., Pedrosa, R., . . . & Mota-Cardoso, R. (2014). Teaching communication skills in clinical settings: Comparing two applications of a comprehensive program with standardized and real patients. *BMC Medical Education*, *14*(1), 92. doi:10.1186/1472-6920-14-92

Chen, D., Lew, R., Hershman, W., & Orlander, J. (2007). A cross-sectional measurement of medical student empathy. *Journal of General Internal Medicine*, *22*(10), 1434-1438.

Clauser, B. E., Margolis, M. J., & Swanson, D. B. (2002). An examination of the contribution of computer-based case simulations to the USMLE step 3 examination. *Academic Medicine, 77*(10), S80-S82.

Cleland, J. A., Abe, K., & Rethans, J. J. (2009). The use of simulated patients in medical education: AMEE Guide No 42 1. *Medical Teacher*, *31*(6), 477-486.

Colliver, J. A., Conlee, M. J., Verhulst, S. J., & Dorsey, J. K. (2010). Reports of the decline of empathy during medical education are greatly exaggerated: A reexamination of the research. *Academic Medicine*, *85*(4), 588-593.

Cox, M., Irby, D. M., & Epstein, R. M. (2007). Assessment in medical education. *New England Journal of Medicine, 356*(4), 387-396.

Cruess, R. L., & Cruess, S. R. (2008). Expectations and obligations: Professionalism and medicine's social contract with society. *Perspectives in Biology and Medicine, 51*(4), 579-598.

Cruess, R. L., Cruess, S. R., & Steinert, Y. (2009). *Teaching medical professionalism* Cambridge University Press.

Cruess, S. R., Johnston, S., & Cruess, R. L. (2004). "Profession": A working definition for medical educators. *Teaching and Learning in Medicine, 16*(1), 74-76.

Daley, B. J. (1999). Novice to expert: An exploration of how professionals learn. *Adult Education Quarterly, 49*(4), 133-147.

Delvaux, N., Razavi, D., Marchal, S., Bredart, A., Farvacques, C., & Slachmuylder, J. L. (2004). Effects of a 105 hours psychological training program on attitudes, communication skills and occupational stress in oncology: A randomised study. *British Journal of Cancer, 90*(1), 106-114.

Dillon, G. F., Clyman, S. G., Clauser, B. E., & Margolis, M. J. (2002). The introduction of computer-based case simulations into the United States medical licensing examination. *Academic Medicine, 77*(10), S94-S96.

DiMatteo, M. R. (1979). A social-psychological analysis of physician-patient rapport: Toward a science of the art of medicine. *Journal of Social Issues, 35*(1), 12-33.

DiMatteo, M. R., Haskard-Zolnierek, K. B., & Martin, L. R. (2012). Improving patient adherence: A three-factor model to guide practice. *Health Psychology Review, 6*(1), 74-91.

Diseker, R. A., & Michielutte, R. (1981). An analysis of empathy in medical students before and following clinical experience. *Academic Medicine, 56*(12), 1004-1010.

DiTomasso, R. A., & Willard, M. (1991). The development of a patient satisfaction questionnaire in the ambulatory setting. *Family Medicine, 23*(2), 127-131. Retrieved from http://ezproxy.pcom.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-0025973352&site=eds-live&scope=site

Dyche, L. (2007). Interpersonal skill in medicine: The essential partner of verbal communication. *Journal of General Internal Medicine, 22*, 1035-1039.

Dyche, L., & Swiderski, D. (2005). The effect of physician solicitation approaches on ability to identify patient concerns. *Journal of General Internal Medicine, 20*(3), 267-270.

Eisenthal, S., Emery, R., Lazare, A., & Udin, H. (1979). 'Adherence' and the negotiated approach to patienthood. *Archives of General Psychiatry, 36*(4), 393-398.

Elnicki, D. M., Lescisin, D. A., & Case, S. (2002). Improving the National Board of Medical Examiners internal medicine subject exam for use in clerkship evaluation. *Journal of General Internal Medicine, 17*(6), 435-440.

Epstein, R. M., Alper, B. S., & Quill, T. E. (2004). Communicating evidence for participatory decision making. *JAMA: The Journal of the American Medical Association, 291*(19), 2359-2366.

Epstein, R. M., Dannefer, E. F., Nofziger, A. C., Hansen, J. T., Schultz, S. H., Jospe, N., & Henson, L. C. (2004). Comprehensive assessment of professional competence: The Rochester experiment. *Teaching and Learning in Medicine, 16*(2), 186-196.

Epstein, R. M., Franks, P., Fiscella, K., Shields, C. G., Meldrum, S. C., Kravitz, R. L., & Duberstein, P. R. (2005). Measuring patient-centered communication in patient–physician consultations: Theoretical and practical issues. *Social Science & Medicine, 61*(7), 1516-1528.

Epstein, R. M., Franks, P., Shields, C. G., Meldrum, S. C., Miller, K. N., Campbell, T. L., & Fiscella, K. (2005). Patient-centered communication and diagnostic testing. *The Annals of Family Medicine, 3*(5), 415-421.

Epstein, R. M., Hadee, T., Carroll, J., Meldrum, S. C., Lardner, J., & Shields, C. G. (2007). Could this be something serious? *Journal of general internal medicine, 22*(12), 1731-1739.

Epstein, R. M., & Hundert, E. M. (2002). Defining and assessing professional competence. *JAMA: The Journal of the American Medical Association, 287*(2), 226-235.

Feighny, K., Monaco, M., & Arnold, L. (1995). Empathy training to improve physician-patient communication skills. *Academic Medicine, 70*(5), 435-436.

Ferreira, P. H., Ferreira, M. L., Maher, C. G., Refshauge, K. M., Latimer, J., & Adams, R. D. (2013). The therapeutic alliance between clinicians and patients predicts outcome in chronic low back pain. *Physical Therapy, 93*(4), 470-478.

Field, A. (2009). *Discovering statistics using SPSS (and sex and drugs and rock'n'roll)*. Thousand Oaks, CA: Sage.

Fiscella, K., Franks, P., Srinivasan, M., Kravitz, R. L., & Epstein, R. (2007). Ratings of physician communication by real and standardized patients. *The Annals of Family Medicine*, *5*(2), 151-158.

Franks, P., Fiscella, K., Shields, C. G., Meldrum, S. C., Duberstein, P., Jerant, A. F., . . . Epstein, R. M. (2005). Are patients' ratings of their physicians related to health outcomes? *The Annals of Family Medicine, 3*(3), 229-234.

Fuertes, J. N., Anand, P., Haggerty, G., Kestenbaum, M., & Rosenblum, G. C. (2015). The physician–patient working alliance and patient psychological attachment, adherence, outcome expectations, and satisfaction in a sample of rheumatology patients. *Behavioral Medicine, 41*(2), 60-68.

Fuertes, J. N., Mislowack, A., Bennett, J., Paul, L., Gilbert, T. C., Fontan, G., & Boylan, L. S. (2007). The physician–patient working alliance. *Patient Education and Counseling, 66*(1), 29-36. doi:10.1016/j.pec.2006.09.013

Gordon-Lubitz, R. J. (2003). Risk communication: Problems of presentation and understanding. *JAMA: The Journal of the American Medical Association, 289*(1), 95.

Hafferty, F. (2004). Toward the operationalization of professionalism: A commentary. *American Journal of Bioethics, 4*(2), 28-31. doi:10.1162/152651604323097718

Hamelin, N. D., Nikolis, A., Armano, J., Harris, P. G., & Brutus, J. P. (2012). Evaluation of factors influencing confidence and trust in the patient-physician relationship: A

survey of patient in a hand clinic. *Chirurgie De La Main, 31*(2), 83-90. doi:10.1016/j.main.2012.01.005

Heisler, M., Bouknight, R. R., Hayward, R. A., Smith, D. M., & Kerr, E. A. (2002). The relative importance of physician communication, participatory decision making, and patient understanding in diabetes self-management. *Journal of General Internal Medicine, 17*(4), 243-252.

Hibbard, J. H., Mahoney, E. R., Stock, R., & Tusler, M. (2007). Do increases in patient activation result in improved self-management behaviors? *Health Services Research, 42*(4), 1443-1463.

Hilton, S. R., & Slotnick, H. B. (2005). Proto-professionalism: How professionalisation occurs across the continuum of medical education. *Medical Education, 39*(1), 58-65.

Hochberg, M. S., Kalet, A., Zabar, S., Kachur, E., Gillespie, C., & Berman, R. S. (2010). Can professionalism be taught? Encouraging evidence. *The American Journal of Surgery, 199*(1), 86-93. doi:10.1016/j.amjsurg.2009.10.002

Hojat, M., Gonnella, J. S., Mangione, S., Nasca, T. J., & Magee, M. (2003). Physician empathy in medical education and practice: Experience with the Jefferson Scale of Physician Empathy. *Seminars in Integrative Medicine, 1*(1) 25-41.

Hojat, M., Gonnella, J. S., Mangione, S., Nasca, T. J., Veloski, J. J., Erdmann, J. B., . . . Magee, M. (2002). Empathy in medical students as related to academic performance, clinical competence and gender. *Medical Education, 36*(6), 522-527.

Hojat, M., Louis, D. Z., Markham, F. W., Wender, R., Rabinowitz, C., & Gonnella, J. S. (2011). Physicians' empathy and clinical outcomes for diabetic patients. *Academic Medicine, 86*(3), 359-364.

Hojat, M., Louis, D. Z., Maxwell, B. S., Gonnella, M. D., & Joseph, S. (2011). The Jefferson Scale of Empathy (JSE): An update. *Population Health Matters (Formerly Health Policy Newsletter), 24*(2), 5.

Hojat, M., Louis, D. Z., Maxwell, K., Markham, F. W., Wender, R. C., & Gonnella, J. S. (2011). A brief instrument to measure patients' overall satisfaction with primary care physicians. *Family Medicine-Kansas City*, *43*(6), 412-417.

Hojat, M., Mangione, S., Nasca, T. J., Gonnella, J. S., & Magee, M. (2005). Empathy scores in medical school and ratings of empathic behavior in residency training 3 years later. *The Journal of social psychology*, *145*(6), 663-672.

Hojat, M., Mangione, S., Nasca, T. J., Ratner, S., Erdmann, J. B., Gonnella, J. S., & Magee, M. (2004). An empirical study of decline in empathy in medical school. *Medical Education, 38*(9), 934-941.

Hojat, M., Spandorfer, J., Louis, D. Z., & Gonnella, J. S. (2011). Empathic and sympathetic orientations toward patient care: Conceptualization, measurement, and psychometrics. *Academic Medicine, 86*(8), 989-995.

Hojat, M., Vergare, M. J., Maxwell, K., Brainard, G., Herrine, S. K., Isenberg, G. A., . . . Gonnella, J. S. (2009). The devil is in the third year: A longitudinal study of erosion of empathy in medical school. *Academic Medicine, 84*(9), 1182-1191.

Hudson, G. R. (1993). Empathy and technology in the coronary care unit. *Intensive and Critical Care Nursing, 9*(1), 55-61.

Imel, Z. E., Baldwin, S. A., Baer, J. S., Hartzler, B., Dunn, C., Rosengren, D. B., & Atkins, D. C. (2014). Evaluating therapist adherence in motivational interviewing by comparing performance with standardized and real patients. *Journal of consulting and clinical psychology, 82*(3), 472-481.

Jackson, J. L., Chamberlin, J., & Kroenke, K. (2001). Predictors of patient satisfaction. *Social Science & Medicine, 52*(4), 609- 620.

Jahng, K. H., Martin, L. R., Golin, C. E., & DiMatteo, M. R. (2005). Preferences for medical collaboration: Patient–physician congruence and patient outcomes. *Patient Education & Counseling, 57*(3), 308-314. doi:10.1016/j.pec.2004.08.006

Janeway, D. (2009). An integrated approach to the diagnosis and treatment of anxiety within the practice of cardiology. *Cardiology in Review, 17*(1), 36-43.

Kimmelman, M., Giacobbe, J., Faden, J., Kumar, G., Pinckney, C. C., & Steer, R. (2012). Empathy in osteopathic medical students: A cross-sectional analysis. *JAOA: Journal of the American Osteopathic Association, 112*(6), 347-355.

Kravitz, R. L., Bell, R. A., Azari, R., Krupat, E., Kelly-Reif, S., & Thom, D. (2002). Request fulfillment in office practice: Antecedents and relationship to outcomes. *Medical Care, 40*(1), 38-51.

Langenau, E., Pugliano, G., & Roberts, W. (2011). Competency-based classification of COMLEX-USA cognitive examination test items. *JAOA: Journal of the American Osteopathic Association, 111*(6), 396-402.

Levinson, W. (1994). Physician-patient communication. *JAMA: The Journal of the American Medical Association, 272*(20), 1619-1620.

Lin, C. T., Albertson, G. A., Schilling, L. M., Cyran, E. M., Anderson, S. N., Ware, L., & Anderson, R. J. (2001). Is patients' perception of time spent with the physician a determinant of ambulatory patient satisfaction? *Archives of Internal Medicine, 161*(11), 1437-1442.

Lockwood, M. D., Tucker-Potter, S., & Sargentini, N. J. (2009). Curricular analysis of competency-based osteopathic medical education: Application of a matrix for quality enhancement to a standardized patient encounter example. *JAOA: Journal of the American Osteopathic Association, 109*(9), 486-500.

Marvel, M. K., Epstein, R. M., Flowers, K., & Beckman, H. B. (1999). Soliciting the patient's agenda. *JAMA: The Journal of the American Medical Association, 281*(3), 283-287.

Miller, S. M., Brody, D. S., & Summerton, J. (1988). Styles of coping with threat: Implications for health. *Journal of Personality and Social Psychology, 54*(1), 142-148.

Moffat, K. J., McConnachie, A., Ross, S., & Morrison, J. M. (2004). First year medical student stress and coping in a problem-based learning medical curriculum. *Medical Education, 38*(5), 482-491.

Moore, P. M., Rivera, M. S., Grez, A. M., & Lawrie, T. A. (2013). Communication skills training for healthcare professionals working with people who have cancer. *The Cochrane database of systematic reviews, 3*, CD003751.

Mueller, P. S. (2009). Incorporating professionalism into medical education: The Mayo Clinic experience. *The Keio Journal of Medicine, 58*(3), 133-143.

National Board of Osteopathic Medical Examiners (2011). *Fundamental osteopathic medical competency domains: Guidelines for osteopathic medical licensure and practice of osteopathic medicine.* Conshohocken, PA: Author.

Newton, B. W., Barber, L., Clardy, J., Cleveland, E., & O'Sullivan, P. (2008). Is there hardening of the heart during medical school? *Academic Medicine, 83*(3), 244-249.

Norcini, J. J., Blank, L. L., Duffy, F. D., & Fortna, G. S. (2003). The mini-CEX: A method for assessing clinical skills. *Annals of Internal Medicine, 138*(6), 476-481.

Norcini, J. J., & McKinley, D. W. (2007). Assessment methods in medical education. *Teaching and Teacher Education, 23*(3), 239-250.

Norman, G. (2002). The long case versus objective structured clinical examinations: The long case is a bit better, if time is equal. *BMJ: British Medical Journal, 324*(7340), 748-749.

Novack, D. H., Epstein, R. M., & Paulsen, R. H. (1999). Toward creating physician-healers: Fostering medical students' self-awareness, personal growth, and well-being. *Academic Medicine*, *74*(5), 516-520.

Oliver, T., Hausdorf, P., Lievens, F., & Conlon, P. (2014). Interpersonal dynamics in assessment center exercises effects of role player portrayed disposition. *Journal of Management*, doi:10.1177/0149206314525207

Ong, L. M., De Haes, J. C., Hoos, A. M., & Lammes, F. B. (1995). Doctor-patient communication: A review of the literature. *Social Science & Medicine, 40*(7), 903-918.

O'Sullivan, A. J., & Toohey, S. M. (2008). Assessment of professionalism in undergraduate medical students. *Medical Teacher, 30*(3), 280-286. doi:10.1080/01421590701758640

Papadakis, M. A., Teherani, A., Banach, M. A., Knettler, T. R., Ratner, S. L., Stern, D. T., . . . Hodgson, C. S. (2005). Disciplinary action by medical boards and prior behavior in medical school. *New England Journal of Medicine, 353*(25), 2673-2682.

Polanyi, M. (1962). *Personal knowledge: Towards a post-critical philosophy.* London: Psychology Press.

Polanyi, M., & Grene, M. G. (1969). *Knowing and being: Essays.* University of Chicago Press.

Pulito, A. R., Donnelly, M. B., & Plymale, M. (2007). Factors in faculty evaluation of medical students' performance. *Medical Education, 41*(7), 667-675.

Pulito, A. R., Donnelly, M. B., Plymale, M., & Mentzer, Jr, R. M. (2006). What do faculty observe of medical students' clinical performance? *Teaching and learning in medicine, 18*(2), 99-104.

Reinders, M. E., Blankenstein, A. H., van Marwijk, H. W., Knol, D. L., Ram, P., van der Horst, H. E., . . . & van der Vleuten, C. P. (2011). Reliability of consultation skills assessments using standardized versus real patients. *Medical Education, 45*(6), 578-584.

Sandella, J. M., Roberts, W. L., & Langenau, E. E. (2011). Relationship between encounter time and candidate performance on COMLEX-USA level 2-PE. *JAOA: Journal of the American Osteopathic Association, 111*(1), 38-43.

Schoenthaler, A., Montague, E., Baier Manwell, L., Brown, R., Schwartz, M. D., & Linzer, M. (2014). Patient–physician racial/ethnic concordance and blood pressure control: The role of trust and medication adherence. *Ethnicity & Health, 19*(5), 565-578.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action* (Vol. 5126). Basic books.

Schwid, H. A., Rooke, G. A., Carline, J., Steadman, R. H., Murray, W. B., Olympio, M., . . . Wetstone, S. (2002). Evaluation of anesthesia residents using mannequin-based simulation: A multiinstitutional study. *Anesthesiology, 97*(6), 1434-1444.

Shanafelt, T. D., West, C., Zhao, X., Novotny, P., Kolars, J., Habermann, T., & Sloan, J. (2005). Relationship between increased personal well-being and enhanced empathy among internal medicine residents. *Journal of General Internal Medicine, 20*(7), 559-564.

Shirazi, M., Labaf, A., Monjazebi, F., Jalili, M., Mirzazadeh, M., Ponzer, S., & Masiello, I. (2014). Assessing medical students' communication skills by the use of standardized patients: Emphasizing standardized patients' quality assurance. *Academic Psychiatry*, *38*(3), 354-360.

Silver, H. K., & Glicken, A. D. (1990). Medical student abuse. *JAMA: The Journal of the American Medical Association, 263*(4), 527-532.

Stern, N. H. (2006). *Stern Review: The economics of climate change* (Vol. 30). London: HM treasury.

Stimmel, B., Cohen, D., Fallar, R., & Smith, L. (2006). The use of standardised patients to assess clinical competence: does practice make perfect? *Medical Education*, *40*(5), 444-449.

Street, R. L., & Buller, D. B. (1987). Nonverbal response patterns in physician-patient interactions: A functional analysis. *Journal of Nonverbal Behavior, 11*(4), 234-253.

Suchman, A. L., Markakis, K., Beckman, H. B., & Frankel, R. (1997). A model of empathic communication in the medical interview. *Journal of the American Medical Association, 277*, 678-682.

Sweeny, K., Shepperd, J. A., & Han, P. K. (2013). The goals of communicating bad news in health care: do physicians and patients agree? *Health Expectations*, *16*(3), 230-238.

Tamblyn, R., Abrahamowicz, M., Dauphinee, D., Wenghofer, E., Jacques, A., Klass, D., . . . & Hanley, J. A. (2007). Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. *Jama*, *298*(9), 993-1001.

Thom, D. H., Kravitz, R. L., Bell, R. A., Krupat, E., & Azari, R. (2002). Patient trust in the physician: Relationship to patient requests. *Family Practice, 19*(5), 476-483.

Travaline, J. M., Ruchinskas, R., & D'Alonzo, G. E. (2005). Patient-physician communication: Why and how. *JAOA: Journal of the American Osteopathic Association, 105*(1), 13-18.

Van Der Vleuten, C. P. (1996). The assessment of professional competence: Developments, research and practical implications. *Advances in Health Sciences Education, 1*(1), 41-67.

Van Der Vleuten, C., Norman, G., & Graaff, E. (1991). Pitfalls in the pursuit of objectivity: Issues of reliability. *Medical Education, 25*(2), 110-118.

Van Zanten, M., Boulet, J. R., Norcini, J. J., & McKinley, D. (2005). Using a standardised patient assessment to measure professional attributes. *Medical Education, 39*(1), 20-29.

VandeKieft, G. K. (2001). Breaking bad news. *American Family Physician, 64*(12), 1975-1978.

Veloski, J., & Hojat, M. (2006). Measuring specific elements of professionalism: Empathy, teamwork, and lifelong learning. In S. T. David (Ed.), *Measuring medical professionalism* (pp. 117–145). Oxford: Oxford University Press

Wass, V., Jones, R., & Van der Vleuten, C. (2001). Standardized or real patients to test clinical competence? the long case revisited. *Medical Education, 35*(4), 321-325.

Weidner, A. C., Gimpel, J. R., Boulet, J. R., & Solomon, M. (2010). Using standardized patients to assess the communication skills of graduating physicians for the Comprehensive Osteopathic Medical Licensing Examination (COMLEX) Level 2-Performance Evaluation (Level 2-PE). *Teaching and learning in medicine, 22*(1), 8-15.

West, C. P., & Shanafelt, T. D. (2007). The influence of personal and environmental factors on professionalism in medical education. *BMC Medical Education, 7*(1), 1-9.

Zhang, X., & Roberts, W. L. (2013). Investigation of standardized patient ratings of humanistic competence on a medical licensure examination using Many-Facet Rasch Measurement and generalizability theory. *Advances in Health Sciences Education, 18*(5), 929-944.

Zolnierek, K. B. H., & DiMatteo, M. R. (2009). Physician Communication and Patient Adherence to Treatment: A Meta-analysis. *Medical Care, 47*(8), 826-834.